Juha Kännö

A short-term price forecast model for the Nordic electricity markets

School of Science

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 23.4.2013

Thesis supervisor:

Prof. Ahti Salo

Thesis instructor:

M.Sc. (Econ.) Henri Äijö

Aalto University
School of Science

AALTO UNIVERSITY
SCHOOL OF SCIENCE

ABSTRACT OF THE MASTER'S THESIS

Author: Juha Kännö

Title: A short-term price forecast model for the Nordic electricity markets

Date: 23.4.2013 | Language: English | Number of pages: 59

Department of Mathematics and Systems Analysis

Professorship: Applied mathematics | Code: Mat-2

Supervisor: Prof. Ahti Salo

Instructor: M.Sc. (Econ.) Henri Äijö

The Nordic countries have a common wholesale electricity market, where the price of electricity is determined by principles of supply and demand. The electricity exchange, Nord Pool Spot, calculates each day wholesale, or spot prices for each hour of the day ahead. Price forecasts are used by electricity producers and large consumers to support operational planning and trading decisions.

Short-term price forecasting typically aims at predicting spot prices up to two weeks ahead. Many existing short-term forecasting approaches are based on time-series methods. They are appealing, because forecasting requires at minimum only historical price data, and the model specification does not heavily rely on the structure of the underlying market. However, the utilization of both external variables and structural information could increase the transparency of the forecasting process and enable the analysis of different scenarios.

This thesis presents the main elements of the Nordic electricity market and reviews prior research on electricity price forecasting models. Based on the findings, a framework for short-term price forecasting is proposed. Price is modelled by combining an external demand estimate with supply functions. A heuristic algorithm is developed for the estimation of supply functions from historical market data, and an attempt is made to explain changes in the pricing of supply by fundamental drivers. The implementation of the framework is tested against historical market data from years 2009–2011.

Hourly spot prices can be represented by the proposed framework under most circumstances. Temporal shifts of supply functions representing particular days can be used to analyse the effects of market fundamentals on pricing. It was found that the effects of fundamentals are highly dynamic and challenging to quantify.

Keywords: electricity markets, spot price, forecasting,
fundamental models, supply function

AALTO-YLIOPISTO
PERUSTIETEIDEN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Juha Kännö

Työn nimi: Sähkön hinnan ennustaminen lyhyellä aikavälillä pohjoismaisilla sähkömarkkinoilla

Päivämäärä: 23.4.2013 Kieli: Englanti Sivumäärä: 59

Matematiikan ja systeemianalyysin laitos

Professuuri: Sovellettu matematiikka Koodi: Mat-2

Valvoja: Prof. Ahti Salo

Ohjaaja: KTM Henri Äijö

Pohjoismailla on yhteiset sähkön tukkumarkkinat, joilla sähkön hinta määräytyy kysynnän ja tarjonnan mukaan. Sähköpörssi Nord Pool Spot laskee päivittäin sähkön tukkuhinnat eli niin sanotut spot-hinnat seuraavan vuorokauden jokaiselle tunnille. Sähköntuottajat ja -kuluttajat tarvitsevat hintaennusteita operatiivisen suunnittelun ja pörssikaupankäynnin tueksi.

Lyhyen aikavälin ennustusmallien tavoitteena on tyypillisesti ennustaa tuntikohtaiset spot-hinnat kahden viikon päähän. Lyhyellä aikavälillä spot-hintoja ennustetaan usein aikasarjamalleilla. Ne ovat houkuttelevia, koska syötteeksi riittää vähimmillään historiallinen hintadata, eikä mallin määrittely riipu vahvasti alla olevan markkinan rakenteesta. Hyödyntämällä sekä ulkoisia muuttujia että rakenneinformaatiota voidaan ennustamisesta tehdä läpinäkyvämpää ja analysoida erilaisia skenaarioita.

Tässä diplomityössä tarkastellaan pohjoismaisten sähkömarkkinoiden peruspiirteitä ja aikaisempaa hinnan ennustamiseen liittyvää tutkimusta. Tähän perustuen esitetään malli, jossa spot-hinta lasketaan yhdistämällä ulkoinen kysyntäennuste tarjontafunktioihin. Työssä kehitetään heuristinen algoritmi tarjontafunktioiden estimointiin markkinadatasta. Lisäksi työssä tutkitaan tarjonnan hinnoittelun riippuvuutta markkinafundamenteista. Mallia testataan markkinadatalla vuosilta 2009–2011.

Mallilla pystytään kuvaamaan tuntitason spot-hintoja useimmissa tilanteissa. Päiväkohtaisten tarjontafunktioiden muutoksia vertailemalla voidaan analysoida markkinafundamenttien vaikutusta hinnoitteluun. Fundamenttien vaikutukset ovat hyvin dynaamisia ja niiden kvantifiointi on haasteellista.

Avainsanat: sähkömarkkinat, spot-hinta, ennustaminen, fundamenttipohjaiset mallit, tarjontafunktio

Acknowledgements

The completion of this thesis would not have come about without the help and support of several people, to say nothing of the innumerable cups of coffee which fuelled the train of thought.

The thesis was written as an assignment from UPM. First I wish to thank Christian Hoffmann, Director of Energy Market Analysis, for entrusting me with the opportunity to work on this topic. I am grateful to my instructor Henri Äijö, who helped me to develop an understanding of the fundamentals of electricity markets. I would also like to extend my appreciation to the members of the project steering group. I am indebted to Timo Javanainen for subtle guidance and many encouraging discussions, Anne Särkilahti for perceptive comments and Ljubov Fedetskina for feedback in application development. Moreover, special thanks go out to Lari Järvenpää for the great company to toss ideas around. All colleagues, with whom I have had the pleasure to work with, have my sincere gratitude for their time and support.

I want to thank my supervisor, Professor Ahti Salo, for guidance in the thesis development and insightful teaching, which has cultivated my understanding of systems analysis and its application to actual business cases. I am also grateful to Dr. Afzal Siddiqui for suggestions that helped me to improve the thesis towards the end of the writing.

I would like to express my appreciation to the friends who have accompanied me on the road to graduation. It would overshadow the length of the actual thesis to mention individually and equally each wonderful person who deserves it. However, it is thanks to all of you that I have grown to be the person that I am now.

Finally, I want to thank my family for their love, care and support.

Helsinki 28.4.2013

Juha Kännö

Contents

Symbols and abbreviations

Symbols

Calculation of short-run marginal cost

c_{fuel}	Cost of fuel
c_{CO_2}	Cost of emission allowance
k_{FX}	Currency exchange rate
k_{HV}	Heat value of fuel
k_{eff}	Efficiency rate of power plant
k_{CO_2}	Emission factor of power plant

Hydro reservoir equations

R_t	Reservoir level at the end of period t
$\overline{R}_t$	Upper reservoir level
$\underline{R}_t$	Lower reservoir level
w_t	Inflow into the reservoir in period t
e_t^H	Water released from the reservoir in period t

Supply function estimation algorithm

t_i	Time
q_i	Quantity
p_i	Price
$x_i = \{t_i, q_i, p_i\}$	Observation of supply data
$X = \{x_i\}$	Set of supply data observations
$X_{SupplyFunction}$	Set of data points that define the estimated supply function
v_{int}	Volume interval parameter of the estimation algorithm

Fundamental regression models

h	Hour
d	Day
$f_{d,h}$	Actual value of a fundamental
$\bar{f}$	Base value of a fundamental
$x_{d,h}$	Value of the fundamental predictor variable
$p_{d,h}$	Price
$q_{d,h}$	Demand volume
β_i	Regression coefficients
$\varepsilon_{d,h}$	Error term

Abbreviations

ANN	Artificial neural network
ARIMA	Autoregressive integrated moving average (time-series model)
CCGT	Combined cycle gas turbine
CET	Central European Time
CfD	Contract for difference
CHP	Combined heat and power
EMPS	EFI's Multi-area Power-market Simulator
MWh_e	Megawatt-hour of electricity
MWh_{th}	Megawatt-hour of thermal energy
PDP	Price-dependent production
PIP	Price-independent production
SFE	Supply function equilibrium
SRMC	Short-run marginal cost
TSO	Transmission system operator
UMM	Urgent market message

Chapter 1

Introduction

1.1 Background and motivation

With the liberalization of electricity markets, power has become a tradable commodity. As a consequence, formerly regulated electricity tariffs have made way for market-based price determination (Nord Pool Spot, 2011b). Electricity differs inherently from other goods in that it cannot be stored. Hence, the amount of power fed into the grid must constantly match consumption to ensure the stability of the power system. In addition, the transmission of electricity requires an extensive infrastructure. These facts are essential to the nature of the power market (see e.g. Weron, 2007).

Deregulation of the Nordic market started in Norway in the early 1990s (Flatabø et al., 2003). Today Norway, Sweden, Finland, Denmark and Estonia form a single market, where electricity is traded in a common power exchange operated by Nord Pool Spot AS (Nord Pool Spot, 2011b). Its main market place is the day-ahead auction market *Elspot*, which determines the spot price. According to Nord Pool Spot (2011b), the power exchange serves the society first by providing transparent pricing for wholesale electricity trade. Second, the spot price is used as reference price in the electricity derivatives market, and quotes for long-term contracts reflect the expectations for future electricity prices. Third, the market provides a system for maintaining a balance between physical supply and demand.

Producers, retailers, and traders, as well as big end-users of electricity, participate in the day-ahead market. Participants submit bids, which express combinations of how much energy they are willing to buy or sell at a particular price, for each hour of the following day. After the closing of the bidding period, all bids are fed into the price determination mechanism. Sell offers are accepted in increasing price order and buy offers in decreasing price order, until an equilibrium price and volume are met (Nord Pool Spot, 2011c). Producers with flexible generation assets can use a detailed price forecast for price-setting purposes, or to optimize their production schedule in order to maximize profits (Bunn, 2000).

Price forecasting activities can be classified by the time horizon, resolution, and purpose of use (Weron and Misiorek, 2006). It is conventional to talk about long-term, medium-term and short-term price forecasting. Long-term price forecasting supports strategic decisions, such as investments in new power plants, and has typically a time horizon of several years. Medium-term price forecasting is used for balance sheet forecasts and risk management, and typically produces a probability distribution of estimated prices (e.g. Eydeland and Wolyniec, 2003; Bunn and Karakatsani, 2003). The granularity of medium and long-term forecasts spans from weeks to years. Short-term forecasting, on the other hand, attempts to predict actual prices in the day-ahead market. In the case of the Nordic market, the main interest lies in forecasting hourly prices from the following day up to one or two weeks ahead.

Short-term forecasting is challenging due to the complexity of the market. The price of electricity is related to external factors referred to as *market fundamentals*. Even though it is known that—for instance—fuel prices, temperatures, precipitation, and electricity prices of neighbouring markets affect the Nordic prices (e.g. Kalatie, 2005; Vehviläinen and Pyykkönen, 2005; Johnsen, 2001), quantifying their joint effect on the price is difficult. Moreover, human psychology induces seemingly irrational market behaviour (Jabłonska, 2011).

A great deal of research has concentrated on modelling spot prices with time series models (e.g. Nogales et al., 2002; Guthrie and Videbeck, 2007), whose performance can be improved by incorporating fundamental external variables (e.g. Weron and Misiorek, 2008; Jónsson, 2008) or regime-switching properties (e.g. Karakatsani and Bunn, 2008b). Another line of research is based on artificial neural networks (ANN), which are good at capturing complex and non-linear effects (e.g. Gao et al., 2000; Catalão et al., 2007; Livanis and Zapranis, 2007). These approaches can be characterized as 'black-box', as they do not represent the structure of the underlying market. Because these models do not answer to the question, through which mechanism changes in market fundamentals affect the spot price, they have a limited value in the analysis of different fundamental scenarios. Short-term forecasting calls for more transparent approaches. This thesis investigates the usability of a 'grey-box' approach, which incorporates knowledge about the structure of the electricity market into the model itself, and could thus improve the credibility and accuracy of the model.

1.2 Research objectives

The purpose of this thesis is *to develop and evaluate a model for forecasting the hourly prices in Nord Pool Spot's day-ahead market one week ahead.* The model is intended to support support operational planning and trading decisions. From this basis the objectives of the thesis are to

1. identify and analyse key factors in spot price formation
2. develop a modelling framework for forecasting hourly prices
3. implement a forecasting model based on the framework.

Electricity prices are characterized by a periodic profile, spikes and a tendency to revert to a mean level (e.g. Bunn and Karakatsani, 2003; Weron, 2005; Meyer-Brandis and Tankov, 2008). Considering the model's intended purpose of use, the design should provide accurate estimates both in terms of hourly profile and price level, and enable sensitivity analysis with respect to market fundamentals.

The scope of this thesis is limited to the *system price*, which is calculated without setting constraints on the transmission capacities between different zones in the market (Nord Pool Spot, 2011b). In the case that transmission lines are congested, there will be different prices in the affected zones, or so-called *bidding areas*. Modelling the *area prices* adds an extra layer of complexity, which constitutes an interesting topic for further studies.

1.3 Structure of the thesis

The remainder of this thesis is organized as follows. Chapter 2 introduces the main elements of the Nordic power market. Chapter 3 reviews prior research on electricity price modelling and forecasting. Emphasis is put on short-term methods, but other approaches are also presented briefly to provide a general view of the field. Chapter 4 proposes a short-term spot price forecasting framework, which is validated against historical market data. The analysis focuses particularly on the effects of market fundamentals and their implications on practical forecasting activities. Finally, Chapter 5 presents the conclusions of the research.

Chapter 2

The Nordic electricity market

2.1 Consumption

Electricity consumption in the Nordic countries is mainly made up by electricity-intensive industries and direct electricity heating (NordREG, 2012). Largest sources of consumption vary by country, but generally follow temperature variations and economic growth. In 2010, the total electrical energy consumption of Norway, Sweden, Finland and Denmark amounted to 361 TWh (Eurostat, 2012). Industrial use had a 41 % share of total consumption. Paper and pulp are the most consuming industrial sector in Finland and Sweden, while metal and petrochemical industries are the biggest consumers in Norway. Denmark has considerably lower industrial consumption. In total, residential consumption amounted to 31 % of the total consumption, while the share of services was 25 %. The price-elasticity of consumption is very low in the short run (Bye and Hansen, 2008), and found to increase along with the spot price (Johnsen, 2001).

Consumption exhibits strong periodical patterns, which are illustrated in Figure 2.1. First, there is a seasonal trend which follows the temperature due to the widespread use of electrical heating. Second, consumption is lower in the weekends than during working days, because many businesses are inactive during weekends. A similar effect can be observed on public holidays. Third, there is a strong intra-day profile. The morning peak occurs when people arrive to their working places, and the evening peak is related to increased household consumption when people come home from work (NordREG, 2012).

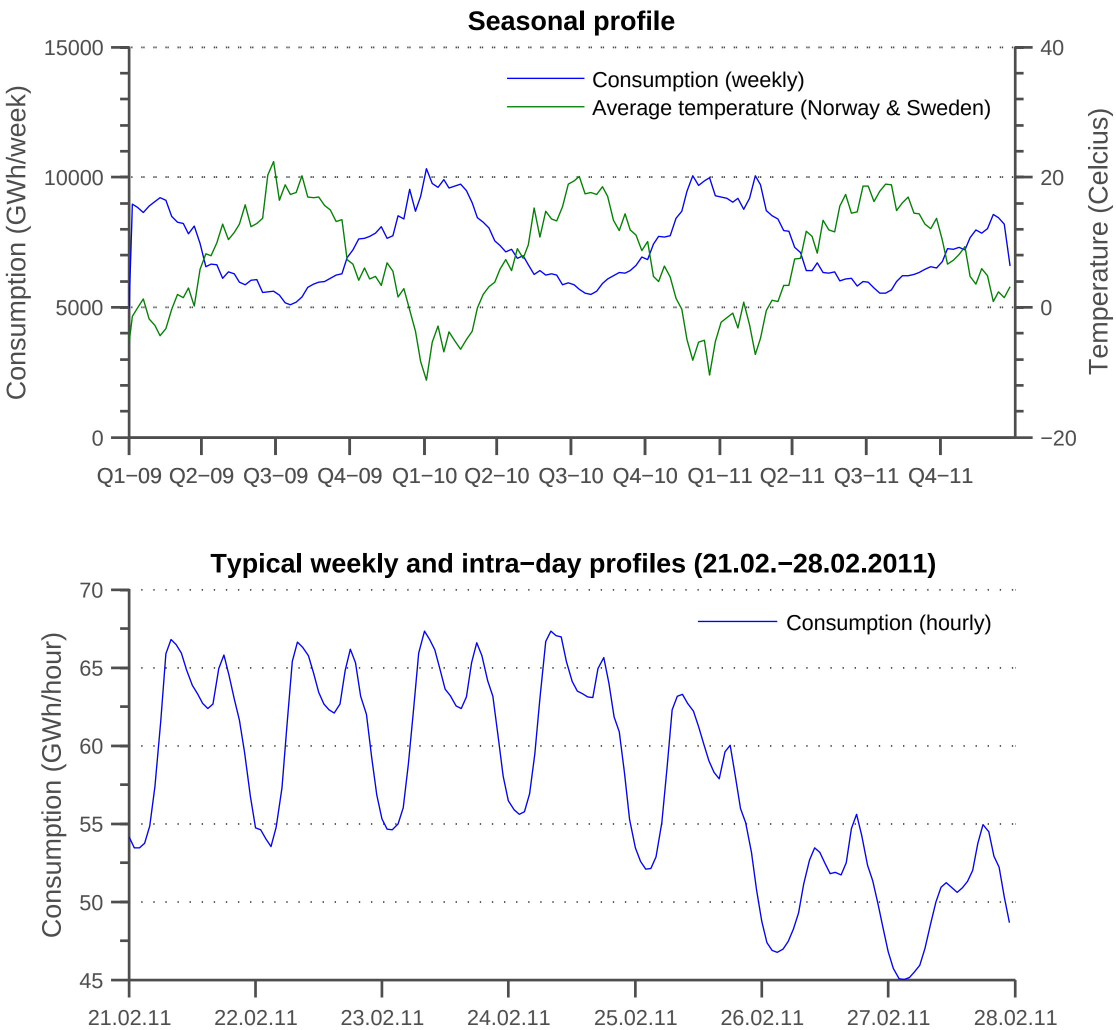


Figure 2.1: Seasonal, weekly, and intra-day profiles of electricity consumption

Table 2.1: Nordic Generation capacity (MW) by power source in 2011 (NordREG, 2012)

	Denmark	Finland	Norway	Sweden	Nordic region
Total installed capacity	13 540	16 713	31 714	36 447	98 414
Nuclear power	—	2 716	—	9 363	12 079
Other thermal power	9 582	10 651	1 062	7 988	29 283
Condensing power	1 590	2 155	—	1 623	5 368
CHP (district heating)	7 118	4 300	—	3 551	14 969
CHP (industry)	674	3 362	—	1 240	5 276
Gas turbines etc.	200	834	—	1 574	2 608
Hydro power	9	3 149	30 140	16 197	49 495
Wind power	3 949	197	512	2 899	7 557

2.2 Generation

Electricity is mainly generated in thermal, nuclear, hydro and wind power plants in the Nordic countries. As a unique feature of the market, around 50–60 % of annual generation originates from hydro power. Furthermore, electricity can be imported or exported to/from other power markets in Central and Western Europe, Baltic countries and Russia. This section describes the characteristics of generation technologies summarized in Table 2.1

2.2.1 Conventional thermal generation

Thermal power plants burn fuel such as coal, natural gas, oil or biomass to produce electricity and/or heat. There is a wide range of plant types which are suitable for different fuels, generation capacities and loads.

Condensing power plants use steam turbines to drive electric generators. Exhaust steam from the turbine is condensed in the process, which gives the name for the plant type. In the Nordic market, condensing plants are mainly used during consumption peaks and as reserve power plants (Kara, 2004, p. 67), while combined heat and power (CHP), hydro and nuclear power are used for baseload production.

Combined heat and power plants are designed to generate both electricity and heat. The co-generation leads to slightly lower efficiency in electricity generation as a trade-off to heat generation. Total efficiency in terms of electricity and heat generation is generally higher than with a condensing power plant. Industrial CHP plants are used to produce steam for heat-requiring processes in pulp and paper mills, and in other industries. CHP plants are also used for district heating in urban areas. The ratio of electricity and heat output depends on the plant design. To operate economically, a CHP plant requires a stable and adequate heat load. It suits well for the Nordic climate with long cold periods, when the demand both for electricity and heat is high (Kara, 2004, p. 75).

In combined-cycle gas turbine (CCGT) technology, a gas-fired combustion turbine drives the electric generator. The heat of gas turbine's exhaust, which otherwise would be wasted, is used to generate steam that drives another generator. CCGT plants producing both electricity and heat can reach efficiency rates of over 90 %.

Gas turbines and diesel engines can be brought on-line very fast, and are typically used for peak power or reserve power.

The short-run marginal cost (SRMC) of thermal generation typically used to determine, whether it is profitable to operate a condensing plant in the short run. It is defined as

$$SRMC\,[\mathrm{EUR/MWh_e}] = c_{fuel} \cdot k_{FX} \cdot k_{HV}^{-1} \cdot k_{eff}^{-1} + c_{CO_2} \cdot k_{CO_2} \cdot k_{eff}^{-1},$$

where c_{fuel} [USD/t] is the cost of fuel, k_{FX} [EUR/USD] the exchange rate, k_{HV} [$\mathrm{MWh_{th}}$/t] the heat value of fuel, k_{eff} [$\mathrm{MWh_e/MWh_{th}}$] the efficiency rate of the power plant, c_{CO_2} [$\mathrm{EUR/t_{CO_2}}$] the cost of emission allowance and k_{CO_2} [$\mathrm{t_{CO_2}/MWh_{th}}$] the emission factor. $\mathrm{MWh_{th}}$ and $\mathrm{MWh_e}$ stand for MWh of thermal energy and electricity, respectively. Condensing plants have also notable start-up costs arising mainly from fuel that must be burned in order to bring the plant up to running state (e.g. Førsund, 2007, pp. 116–117). These costs are substantial and must be also taken into account in operating decisions.

2.2.2 Nuclear power

Nuclear power plants use the energy of fission reaction to boil water, but operate otherwise with same principle as condensing power plants. In Finland and Sweden, nuclear power is used for baseload production. Nuclear plants typically have a very high operating rate.

2.2.3 Wind power

Wind power plants use the kinetic energy of wind to drive an electric generator. Wind power's most significant characteristic is its volatility: the output depends on actual wind speed. A power system cannot thus be built on wind power only. Denmark is world leader with a 21 % share of wind power in its electricity supply (World Wind Energy Association, 2011).

2.2.4 Hydro power

Hydro power plants use the potential energy of water to drive water turbines connected to electric generators. Plants can be characterised by how well production can be regulated. Generation of *run-of-river* plants depends on the natural flow in the waterway where the plant is located. Dams can be used to store some water, which gives better control over the production. A *reservoir* plant maintains a large water storage, where water can accumulate. The size of a reservoir can be equal even up to the production volume of several years. Water is typically discharged through plant's turbines, but it can be also directed to spillways leading it past the plant.

Javanainen (2005) notes that the climate and geography in Norway are very suitable for reservoir hydro power production, whereas Finnish and Swedish hydro power plants have less storage capacity, and are thus more dependent on the natural flows. Figure 2.2 compares weekly inflows and hydro production volumes in Norway and Sweden. It appears that the production volumes in Sweden are more strongly connected to inflows, which supports the argument.

Inflow is a measure of water entering the water system, where a plant is located. Its main factors are precipitation and melting snow. Inflow corresponds to an increase in the *reservoir level*, whereas a discharge decreases it. *Hydro balance* measures the deviation of current water resources to a multi-year average value of that time of the year. In addition to the reservoir content, hydro balance takes into account water contained in the soil and snow which will affect the reservoir.

2.2.5 Optimal generation mix

The load-duration curve, illustrated by Figure 2.3, can be used to determine an optimal mix of generation technologies that are used to satisfy the electricity demand.

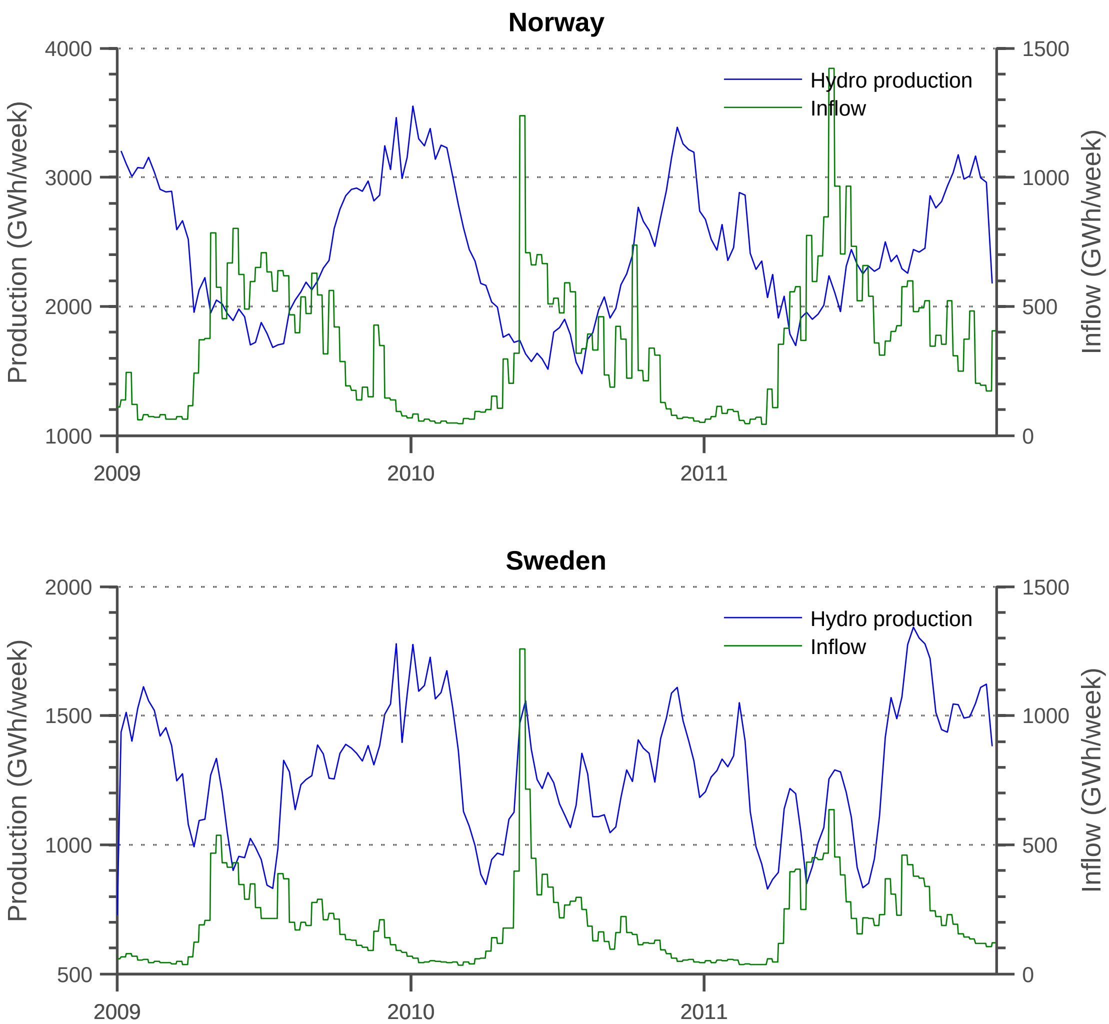


Figure 2.2: Inflow and hydro power production in Norway and Sweden. The timing of inflows correlates between the countries, but Swedish production volumes appear to be more strongly connected to arriving inflows.

The choice of technology depends on the amount of power required and the expected annual operating hours for different load levels. It is economically reasonable to cover baseload with technologies which have low variable costs, while considerable fixed costs will then be distributed to a large number of operating hours. On the other hand, peak load plants and power reserves, which are activated in extreme situations, are typically based on technologies with low fixed costs and higher variable costs, because the fixed costs must be recovered during few annual operating hours. (Vuorinen, 2009)

In the Nordic countries, highest precedence is given to production forms that cannot

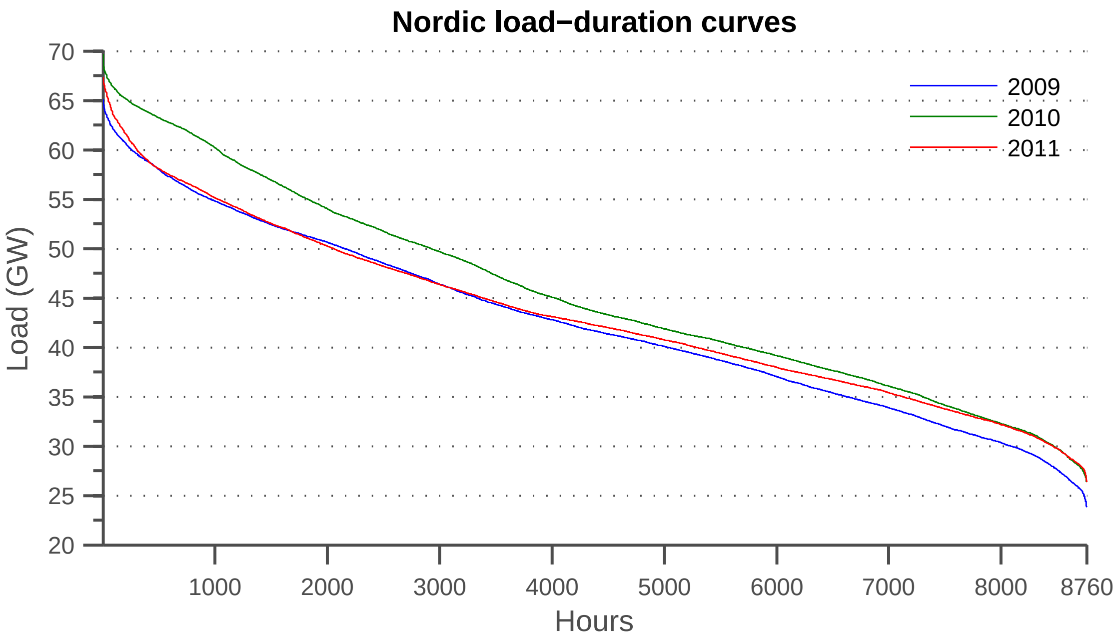


Figure 2.3: Nordic load-duration curves

be regulated, that are wind and a part of hydro power. They are followed by other production forms in an increasing order of marginal production costs, or the so-called *merit order*, as illustrated in Figure 2.4. In addition to wind and unregulated hydro power, baseload production consists of nuclear power and electricity from CHP production. They are followed by condensing power plants, and finally gas turbines and other peak-load plants.

In a perfectly competitive market, the market price will equal the marginal production cost of the most expensive generator that is dispatched. Regulated hydro power is somewhat special in the generation mix: It has extremely low marginal costs, and generation can normally be scaled with high flexibility. Consequently, regulated hydro power is allocated to periods of high demand, when its opportunity cost is high. In the absence of hydro power generation, expensive generators would have to be dispatched in the merit order.

2.3 Power grid

This section outlines the structure and operation of the Nordic power grid. The grid can be broken down to national main grids, regional transmission networks and local distribution networks. Transmission lines between main grids connect the power

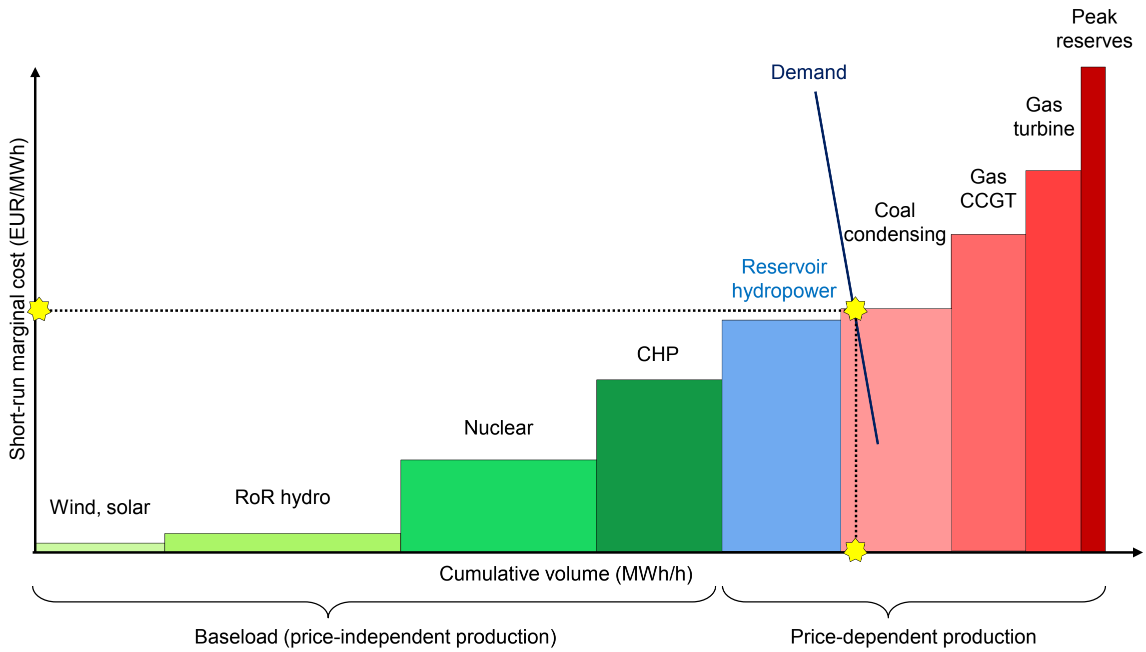


Figure 2.4: Idealised merit order curve of Nordic production. It should be noted that in reality plants using same generation technology can have different marginal production costs.

systems of neighbouring countries. The main grids are owned and maintained by national transmission system operators (TSOs). They are public utilities, which are responsible for the stability of the power system in their area. The Nordic TSOs are Stattnet (Norway), Svenska Kraftnät (Sweden), Fingrid (Finland) and Energinet.dk (Denmark).

One can picture the Nordic market as a system of water tanks, where the water flow represents electric power, and pipes connecting the tanks are transmission lines (Figure 2.5). Producers pour power into the system, and consumers tap it. When the water surface stays at a constant level, demand and supply are in balance. Producers and consumers are not geographically evenly distributed. For instance, most hydro power plants are scattered around Norway, and the northern parts of Sweden and Finland. Due to the limited capacity of transmission lines, arbitrary supply and demand cannot be matched. Instead, local generators will be dispatched in the vicinity of the consumer.

The market uses *bidding areas*[1] as an economic tool to deal with the transmission bottlenecks. Each country makes up at least one bidding area, but TSOs may divide their country into more areas. Each area has an adequate transmission and

[1]The term *price area* used by some sources is interchangeable with *bidding area*.

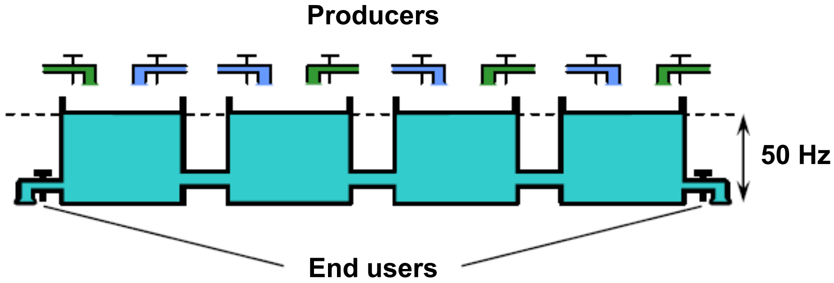


Figure 2.5: Depiction of the power system as system of water tanks. When supply and demand are in balance, the electric current in the grid is at nominal 50 Hz frequency. Imbalances cause voltage and frequency fluctuations, which endanger the stability of the power system. Source: Nord Pool Spot (2011b)

distribution capacity within itself, but transmission lines between areas can become congested. Should it happen that power cannot 'flow' between two bidding areas to match total supply and demand, Nord Pool Spot increases the electricity price in the area with a supply deficit by an amount which lowers the demand to match the available supply. Consequently, the electricity price in the area with supply surplus enjoys a lower price. Since the beginning of 2012 Norway is divided into five bidding areas, Sweden into four areas and Denmark into two areas. Finland and Estonia make up their own bidding areas.

2.4 Electricity trading

This section presents the markets where physical electricity and electricity derivatives are traded. Trading in the physical market involves always a physical delivery of power. The physical market covers day-ahead, intra-day and regulating power markets. The financial markets deal with electricity derivatives, which are settled by cash payments only and do not involve any physical delivery of electricity.

2.4.1 Physical market

Elspot is the main marketplace in Nord Pool Spot. Approximately 74 % of electricity consumed in all Nordic countries was traded in the exchange in 2011 (Nord Pool Spot, 2011a). Elspot is a *day-ahead market*, meaning that every day the power delivery for each hour of the following day is subject to trade. Market participants submit bids where they state the price and quantity of power which they are willing to sell or buy at every hour. Several price steps can be specified. Elspot is a closed auction, and participants have no knowledge about other's bids. Bids are submitted each day until the deadline called *gate closure* at 12:00 CET.

After gate closure the exchange determines spot prices for each hour. The *system price* is a reference price which does not take into account transmission constraints between bidding areas. If available transmission capacities do not constrain power delivery, system price will be the actual price in each bidding area. Otherwise *area prices* are calculated so that demand and supply in each bidding area do not violate transmission constraints. In effect, the area price will be below system price in surplus areas and above it in deficit areas. The prices are normally published between 12:30 to 12:45 CET.

Elbas market supplements Elspot by enabling trading up to one hour before the delivery. There are 12 to 36 hours between gate closure and the delivery hour. Sellers and buyers may need to deviate from the commitments made in Elspot bids. For instance, a drop in temperature could increase demand, or a seller might be unable to produce power due to a plant outage. Elbas quotes the highest buy price and lowest sell price for each hour, and orders are executed immediately when prices match.

Management of the power balance requires still finer market instruments than Elspot and Elbas markets. For this reason the TSOs run in each country a *regulating power* market, where participants can make offers to adjust their generation or consumption capacity within an hour. TSOs can accept these offers in the situation, where consumption exceeds generation (known as *up-regulation*) or generation exceeds consumption (*down-regulation*).

Elspot bid types and the price calculation principle

Willingness to buy or sell power in Elspot can be expressed by different types of bids. *Hourly bid* is the most common type. It consists of a set of price limits and corresponding energy volumes for the applicable delivery hour. An hourly bid must at least specify volumes for the minimum and maximum price limits set by Nord Pool Spot. Values between the given price steps are determined by means of linear interpolation. *Flexible hourly bids* let a participant to sell, but not to buy, energy in any hour of the trading day. The bid can be activated for one hour of the trading day. *Block bids* express willingness to buy or sell power during a minimum time of three consecutive hours. A block bid consists of a price limit, the hourly energy volume and the start and stop times. (Nord Pool Spot, 2011c)

The objective of Nord Pool Spot's price calculation algorithm is to maximize *total social welfare*, which is the sum of consumers' and producers' surplus over all hours, bidding areas and bid types. The problem is subject to several constraints, which ensure that (Nord Pool Spot, 2011c)

1. the volume of purchases and power flow in is equal to the volume of sales and power flow out in each bidding area
2. imports and exports to/from other markets satisfy specified conditions
3. flows between bidding areas do not exceed available transmission capacities
4. flexible hourly bids and block bids are activated only if they increase the total welfare.

Following gate closure, the exchange determines the system price and area prices. In the system price calculation, aggregated supply and demand curves are constructed for each hour from the hourly bids of all bidding areas. The intersection of the curves gives an equilibrium price and turnover. Block bids and flexible hourly bids are activated if their inclusion improves the value of the objective function. Constraint 3 is not used while the system price is calculated. If constraint 3 is binding when in use, bidding areas will have different area prices. In the area price calculation, aggregated supply and demand curves are created for each area from the bids of the market players located in that area. If a bidding area has power surplus (extra power flows to an adjacent area), a volume corresponding to the transmission capacity is added as price-independent demand to the surplus area. The same volume is added as price-independent supply to the deficit area.

2.4.2 Financial market

The spot prices are highly volatile due to the instantaneous nature of the market. Market participants, who are directly exposed to the spot price, face a great uncertainty about their future income or expenditure. A producer may want to fix the sales price for part of their future production, which secures a certain future cash flow even if the price level would drop. Likewise, a big consumer may want to fix the purchase price for a certain volume. Products of the financial market are mainly used for protection against the price risk and proprietary trading.

Standardized contracts are traded in an exchange operated by NASDAQ OMX Commodities. Main products are futures, forwards and options. There are also contracts for differences (CfDs), which can be used to hedge against the difference between the system price and an area price. Financial contracts do not involve any physical delivery of power. They are settled by cash payments which depend on the contract price and the reference price, which is the system price (or in case of CfD, the difference of system and area price) at the time of delivery. Trading financial contracts outside of the exchange is called the *over-the-counter* market. There parties can enter into any type of contracts desired. A bilateral power contract can for example couple price with outside temperature.

The time horizon in the financial market is six years ahead. Within the horizon, futures are available for delivery periods of days and weeks, and forwards for months, quarters and years. A forward curve quotes the prices of contracts with different maturities and thus tells how the market values electricity that will be delivered in the future. Sometimes the forward price includes a risk premium and may thus be very different than a spot price forecast.

2.5 Generation scheduling and pricing

This section discusses generation scheduling activities in a deregulated market, such as the Nordic market, with both hydro and thermal generation capacity. The task of finding an optimal generation schedule is closely related to electricity price forecasting. Depending on the type of scheduling approach, price can be endogenous to the model, or come from an external price forecasting model. This section introduces the concept of *water value*, which will play an important role in the forecasting framework proposed in this thesis.

In regulated markets, the objective of generation scheduling for a utility was to minimize overall costs while satisfying the demand in their concessionary area (Wolfgang et al., 2009). The utilities were also collectively responsible for maintaining power reserves, which were required to ensure the stability of the system. Following the deregulation and the establishment of the power exchange, power producers have turned into market players, and their objective has changed to maximizing profits. Importantly, Wallace and Fleten (2003) point out that this does not change the pricing if producers are required to be efficient: in both cases the market price should follow the marginal cost of production. However, producers are no longer required to use their own generation assets to meet customer obligations, but they can purchase power from the market. Similarly, there are also markets for reserve power, which are maintained by the TSOs.

The planning activities are typically divided into three time horizons, which employ different types of models. Long-term planning is done up to 15–20 years ahead and concerns investment decisions. Medium-term planning has a 1–3 year range and mainly deals with hydro reservoir management. Short-term planning typically covers a time period up to one or two weeks ahead, and deals with the economic dispatch of generating units. Models are typically connected so that the results of a longer-term model are used to set the boundary conditions of a shorter-term model (Wallace and Fleten, 2003).

The principle of hydro power production planning is to maximize value creation by ensuring that as much water as possible is available in high-price periods (e.g. Fosso et al., 1999; Førsund, 2007, p. 116). Long-term planning utilizes commonly models, which attempt to capture dynamics of the whole power system. Examples of such are MARKAL (Seebregts et al., 2001), BALMOREL (Ravn, 2001) and EMPS (Wolfgang et al., 2009). EMPS was developed by the Norwegian research organisation SINTEF and is designed for markets with a large share of hydro production, such as the Nordic market. The model produces an optimal generation schedule that is based on stochastic input variables and information about the hydro power and thermal generation capacities. It is used by many large power producers, as well as in power system studies.

The most critical constraints in the generation scheduling problem are the upper and lower hydro reservoir levels, and coupling the reservoir levels of consecutive periods. When all variables are converted into energy units, these constraints can be stated as

$$R_t \leq R_{t-1} + w_t - e_t^H \tag{2.1}$$

$$\underline{R}_t \leq R_t \leq \overline{R}_t \quad \forall t, \tag{2.2}$$

where R_t is the reservoir level at the end of period t, w_t the inflow into the reservoir in period t and e_t^H the water released from the reservoir in period t. $\underline{R}_t$ and $\overline{R}_t$ are the lower and upper reservoir levels. If no water is lost because of overflow from reservoirs, the total amount of production in period t equals then e_t^H plus generation from non-storable (run-of-river) inflows (Førsund, 2007, pp. 35–38).

In a hydro-thermal market, the immediate opportunity value of water reflects the SRMC of thermal generation that is needed to substitute hydro power. However, because water can be stored to some extent in reservoirs, it is more truthful to base the water value on the expected SRMC of thermal generation and water storage possibilities in several future periods (e.g. Bye and Hansen, 2008). EMPS determines the water values with a stochastic dynamic programming method[2] (Wolfgang et al., 2009). The logic is that *water from a reservoir should be used in the current period if the income is greater than the water value, or the expected marginal value of passing it to the next period.* The water value approaches zero as the reservoir level comes close to its upper limit. Consequently, run-of-river plants and reservoirs with little control have very low water values.

2.6 Fundamentals and characteristics of electricity prices

This section presents fundamental drivers and characteristics of electricity spot prices. Because electricity cannot be stored, it is likely that the price of electricity is driven by market fundamentals behind spot supply and demand more directly than any other commodity (Geman and Roncoroni, 2006). Due to the high share of hydro power generation in the Nordic market, fundaments behind water values have a great impact. Key drivers are the actual and expected values of hydrological fundamentals (inflow, hydro balance, reservoir levels). For reference, see e.g. Botterud et al. (2010), Bye et al. (2006) or Johnsen (2001). Furthermore, the SRMC of condensing production affects directly the price at which it is offered to the market, and is an essential in the determination of water values (Section 2.5). Finally, price is driven by consumption, which is in the Nordic market strongly dependent on outside temperature (Section 2.1).

Seasonality, spikes and mean-reversion are three characteristic features electricity prices used commonly used stochastic modelling. For reference, see e.g. Skantze et al. (2000), Geman and Roncoroni (2006) and Huisman et al. (2007), or Bunn and

[2]The scheduling problem is stochastic due to the uncertainty associated with future inflows and demand, and dynamic because of (2.1), which links together production decisions in each period.

Karakatsani (2003) for an overview. Stochastic models aim at defining a random process which has the statistical properties of actual prices. It should be considered that the division of actual prices to different components is an arbitrary decision. Notwithstanding, this section continues with a discussion of these features in relation to the previously mentioned fundamentals and the market structure.

Seasonality. As illustrated in Figure 2.6, the hourly price profile seems to follow closely the hourly consumption profile. Similarly to consumption, the price level is lower in weekends than during working days. In contrast, the comparison of daily average prices and consumption points out that prices do not equally closely follow the seasonality of consumption in the long run.

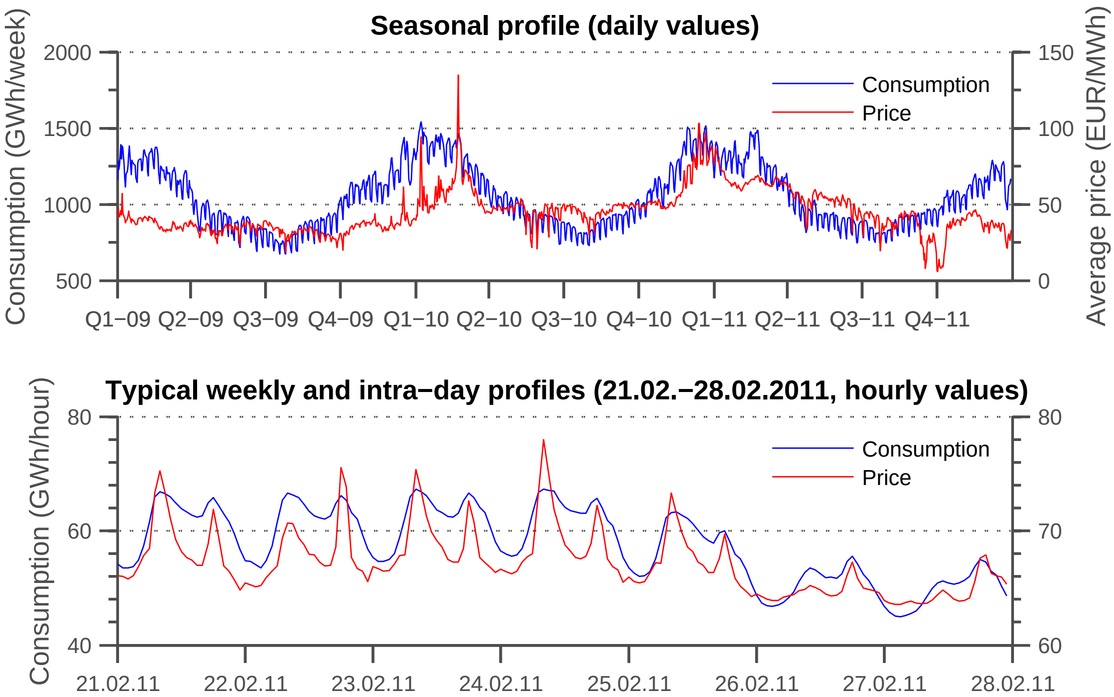


Figure 2.6: Seasonal and intra-day consumption profiles compared to price profiles

Spikes. Given the non-storability of electricity and short-term inelasticity of demand, spikes are often caused by generation outages or transmission failures (e.g. Weron, 2005; Bunn and Karakatsani, 2003). For instance, Vehviläinen et al. (2010) attributed peaks in the Nordic spot price in the winter 2009–2010 to problems with Swedish nuclear supply (Figure 2.7). The underlying reason is that the capacity of electricity supply in the market can be considered fixed in the short term. As demand approaches the total capacity, highly-priced reserve power becomes activated in the merit order (e.g. Kanamura and Ōhashi, 2007). Figure 2.8 illustrates the 'hockey stick' shape of Elspot supply curves, which makes the price extremely sensitive to demand in the vicinity of the capacity limit.

Mean-reversion. Stochastic models (e.g. Skantze et al., 2000; Geman and Roncoroni, 2006; Huisman et al., 2007; Bunn and Karakatsani, 2003) commonly assume that electricity prices revert to some trend line after jumps or spikes. The trend may be assumed to represent seasonal variation, or as in the case of oil, coal and natural gas, the long-term marginal cost of production (Pindyck, 1999). In any case, the trend line itself cannot be directly observed.

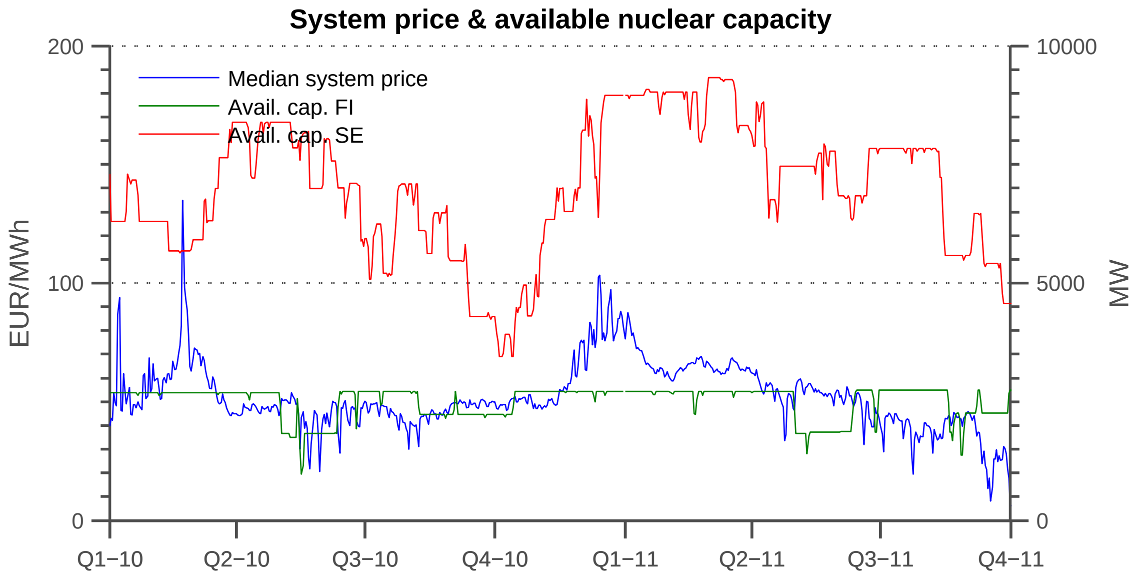


Figure 2.7: System price and available nuclear generation capacity

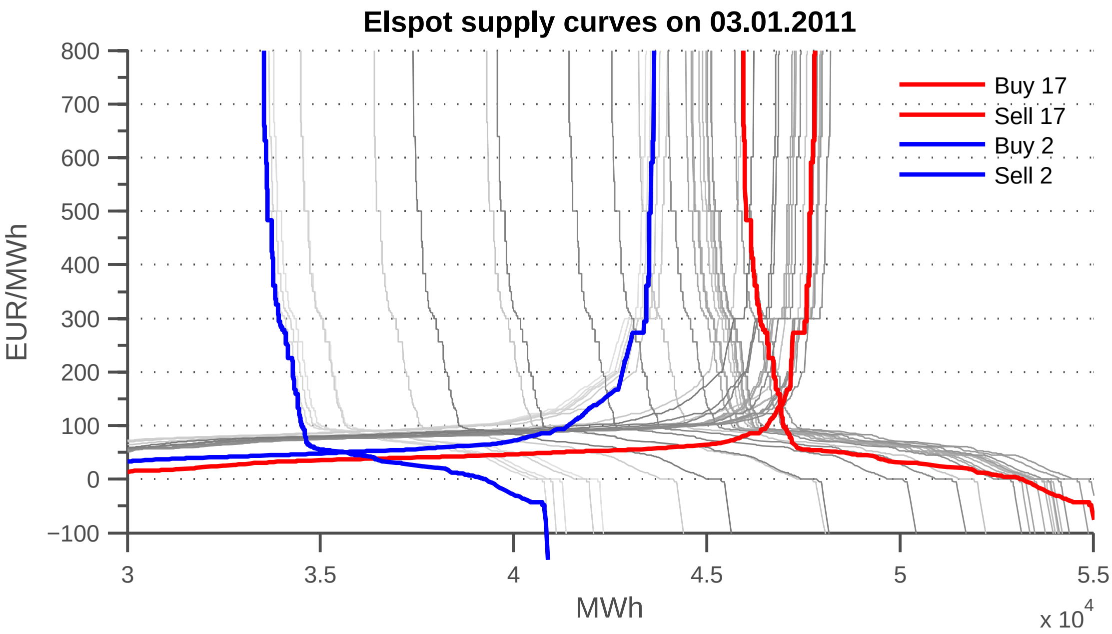


Figure 2.8: Elspot supply curves illustrate the 'hockey stick' shape. 3.1.2011 was the most expensive day of the year. Hourly supply and demand curves are in grey. Lowest price was on hour 2–3 (blue) and highest price on hour 17–18 (red).

Chapter 3

Literature on electricity price models

This chapter reviews previous research on electricity price forecasting with the aim of finding techniques that could be applied in this thesis. The models are sorted into six broadly defined categories, which were adapted from Weron and Misiorek (2006). Each category is discussed in its respective section. Most relevant references to short-term spot price forecasting were found in autoregressive (Section 3.4), artificial intelligence-based (Section 3.5) and fundamental model categories (Section 3.6). Together with the remaining categories, the literature review provides an all-round view of modelling techniques and their applications.

3.1 Cost-based models

Cost-based models attempt to match the estimated demand with supply at minimum cost. Angelus (2001) points out that such approaches were successful in regulated markets with a stable structure, publicly available market information, and planning and coordination between neighbouring utilities. He argues that demand was estimated by scaling historical values, and supply was represented by stacking up the capacities of generating units in the increasing order of their variable operating costs. A price forecast could then be made up by matching regional demand to regional supply, and accurate results could be obtained even on hourly level. However, Angelus (2001) notes the cost-based models are badly suited for deregulated markets, because they were not designed to capture evolving market conditions, uncertainty or market power.

3.2 Game-theoretic models

Game theory studies decision problems, where each player must consider the decisions of other players while making their own decision (Gibbons, 1992). For example, producers' choices of production quantities in an oligopolistic power market constitute such a problem. Kumar David and Wen (2001) note that models of this class are mainly used to simulate outcomes of different market policies, and to look for evidence of market power in existing markets. Two commonly used approaches are *Cournot equilibrium* and *supply function equilibrium* (SFE).

In the Cournot equilibrium, players choose simultaneously their production quantities. In the equilibrium none of the players have an incentive to modify their production quantity. The equilibrium market price is determined by the aggregated demand curve and the total production quantity (Kumar David and Wen, 2001). Early studies of the potential use of market power include Andersson and Bergman (1995) for the Nordic market, and Borenstein et al. (1999) for California.

In the supply function equilibrium, each player chooses a supply function which states the production quantity with regard to the price level (cf. supply quantity in Cournot equilibrium). At the equilibrium none of the players have an incentive to modify their supply function. The equilibrium market price is determined by aggregated supply and demand functions. Kumar David and Wen (2001) argue that SFE offers a more realistic view of electricity markets because suppliers can state their offers both in terms of quantity and price, as opposed to only quantity in the Cournot model. Effects of market power in the UK market have been studied e.g. by Green and Newbery (1992) and Baldick et al. (2004).

According to Bunn and Oliveira (2001), the repetition of a price auction creates an opportunity for market players to experiment with bids and learn from the outcomes. As result, the equilibrium price may shift. They study the behaviour with *an agent-based simulation model*, where sellers and buyers are represented by algorithmic agents acting like conceptual market players. Bunn and Oliveira (2001) conclude that agent-based simulations are best suited for long-term assessments of markets, can be especially useful in predicting the behaviour of a market that does not yet exist in reality.

3.3 Stochastic models

Stochastic models attempt to replicate the statistical properties of electricity prices, and they are typically used to estimate the distribution of future prices. The motivation behind this type of modelling is not forecasting price levels, but rather managing risks carried by the inherent uncertainty in price forecasts (Eydeland and Wolyniec, 2003). The models can focus on spot or forward price. Spot price models account for the characteristic features of prices, which are high volatility, seasonality, occurrence of spikes and reversion to a mean price level (Weron, 2005). Forward price models introduce the risk premium to underlying spot price, and possibly also the number of tradable contracts at each point of time. For reference, see e.g. Meyer-Brandis and Tankov (2008) for an overview of reduced-form spot models, and Bunn and Karakatsani (2003) and Eydeland and Wolyniec (2003) for both spot and forward prices models.

Stochastic models can have structure that resembles fundamental price drivers, as in the stochastic bid model presented by Skantze et al. (2000). In their approach, demand is a stochastic process, and the supply is approximated with an exponential function. The shape of the supply function is fixed, but its temporal shifts are modelled as a stochastic process. The model was calibrated with price and turnover volume data from the market operated by ISO New England. Skantze et al. (2000) relate the shifts in supply to four drivers:

1. Fuel price: An increase in fuel prices increases production costs, and suppliers must ask higher prices in order to stay profitable.
2. Unit outages and scheduled maintenance: Change in the availability of production capacity causes shifts. The size, duration and frequency of their occurrence depend on the technology.
3. Gaming and strategic bidding: Producers with a significant market share may raise the market price by intentionally withdrawing some of their supply.
4. Unit commitment decisions: Generators are subject to constraints and costs related to starting up and shutting down units, which causes generators to bid differently from marginal production costs.

3.4 Autoregressive models

Autoregressive models attempt to describe the behaviour of a variable, such as the spot price, in terms of its own past values. The models are widely used in econometrics and have also a track record in the field of power markets. Variations of autoregressive techniques are used both for consumption and price models. They are attractive, because they are good at capturing seasonal effects, and do not require information about the structure of the underlying market.

The standard technique is based on autoregressive integrated moving average (ARIMA) models (Box and Jenkins, 1976). It assumes that the forecasted variable can be expressed as a linear function of its past values and random noise at each point of time. Consequently, it is assumed that the variable is stationary, meaning that its statistical properties such as mean and variance are constant. The integrated (I) part refers to differencing the variable in order to remove non-stationarity. The autoregressive (AR) part of order p can be written as

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \varepsilon_t,$$

where ϕ_i are parameters, c is a constant and ε_t is a noise term with a mean of zero and constant variance. Similarly, the moving average (MA) part of order q is

$$Y_t = \sum_{j=1}^{q} \theta_j \varepsilon_{t-j},$$

where θ_j are parameters and ε_{t-j} are the values of the noise terms. Summing up AR and MA components gives a full ARMA or ARIMA model specification.

Weron and Misiorek (2006) argue that electricity prices present non-linear dynamics which violate the stationarity assumption, and the problem should be addressed with more advanced tools. One such tool is the Generalized AutoRegressive Conditional Heteroskedastic (GARCH) model. Garcia et al. (2005) find that their GARCH model outperforms a general ARIMA model in day-ahead price forecasting, when prices are highly volatile and spikes occur.

In the Elspot market, participants submit at once bids for each hour of the next day, and the exchange determines and publishes the prices for those hours. Huisman et al. (2007) point out a difference in modelling hourly prices versus daily average prices. Time series models assume that the information set is updated when moving from one observation to the next in time. In a time series of hourly prices, however, each hour of a day has the same information set which is updated over the days. Therefore, direct application of time series models to hourly price data does not have a sound theoretical background. To overcome this issue, they propose a cross-sectional panel framework, where each of the 24 hours of a day is modelled as a separate stochastic process. For example, the price of hour 13 over consecutive days is a time series process, as the information set is updated when moving from one observation to the next. Their results from power markets in the Netherlands, Germany and France show that hours have different mean price levels. Prices in peak hours correlate with each other, and the same applies to off-peak hours.

The forecasting accuracy of time series models can be improved by adding external variables to the model. Several papers identify electricity demand as the main fundamental driver in time series models (e.g. Bunn, 2000; Nogales et al., 2002; Weron and Misiorek, 2008). Jónsson (2008) presents a time series model to forecast the hourly price in the West-Denmark area. The relative share of wind power in total power production is used as an external variable.

Parameters of a time series model, which is intended for short-term forecasting purposes, change with respect to time due to different market situations. (Jónsson, 2008) discusses various weighting methods can be used to construct parameter estimates that evolve consistently in time. Karakatsani and Bunn (2008a) construct a time series model, where the parameters are assumed to follow a random walk process.

3.5 Artificial intelligence-based models

In the scope of time series forecasting, artificial intelligence deals with modelling techniques that take no a priori assumptions about the parameters of the input data, and adapt their internal structure to a data sample through a training process. Artificial neural networks (ANN) are inspired by the structure and functionality of biological neural networks. An ANN consists of interconnected artificial neurons, which process input data. Each neuron can be connected to other neurons or produce output data. In the training typically the weights of neurons are adjusted so that the ANN produces a desired output with a given input. This structure enables ANN's to handle non-linear relationships. A comprehensive foundation on ANN's can be found in e.g. in Haykin (1994).

Bunn (2000) notes that ANN's have proven to be well-suited for electricity load forecasting purposes. They have also been applied directly to prices. Szkuta et al. (1999) analyse the Victorian market in Australia, Livanis and Zapranis (2007) study average daily prices in the Nordic market and Queiroz et al. (2007) look at the Brazilian market. Gao et al. (2000) and Catalão et al. (2007) examine the Californian market. The performance of ANN forecasts is typically compared to conventional linear regression models. The sentiment of these papers is that price forecasting with ANN models is not yet quite mature.

3.6 Fundamental models

Fundamental models describe the electricity price in terms of physical and/or economic variables. The functional relationships of input variables incorporate information about the structure of the market. The values of fundamental variables are typically outputs from other models, such as ones for demand, production cost or hydrology.

In general, the fundamental price forecasting approach is to determine the intersection of demand and supply functions at each time interval in the market. Demand and supply can be modelled with most suitable techniques. Demand forecasts are based on consumption forecasting methods, which typically utilize weather variables but assume no price elasticity. The supply function reflects marginal production costs in a competitive market. Hence, it can be estimated by a merit-order curve of available production capacity. If market concentration is high, game-theoretic approaches may yield more realistic results. However, their applicability in the short-term is limited because of simplifying assumptions, which need to be done in order that the model can be solved. (Bunn, 2000)

Considering the Nordic market, the main challenge of fundamental modelling lies on the supply side. As pointed out in Section 2.5, the pricing of hydro power depends on water values, and their variation affects the shape of the supply function for the whole the Nordic market. In contrast, the short-run marginal cost of thermal generation follows mainly fuel prices. For this reason, models developed for markets with mainly thermal generation may not be applicable.

Dueholm and Ravn (2004) present models of hourly supply functions for Norwegian electricity production, which is hydro power from large reservoirs. They suggest that the supply price is a function of the volume of regulated hydro power production. The model is applied to Norwegian price and production volume data. Because only total hydro power production volumes were publically available, their model attempted to estimate the volumes of regulated and unregulated production. It should be noted that the model does not use inflows or other fundamental variables.

Also Javanainen (2005) studies the Norwegian electricity supply, and finds evidence on strong price-dependency of production. He attributes it to the flexibility of the production system. According to him, the degree of flexibility varies with the time of year, because it is linked to reservoir levels which are affected by seasonal weather patterns. According to Javanainen (2005), the supply function can be characterized by dividing it to three parts representing price-independent production, price-dependent production and maximum production capacity. He notes that inflow correlates with price-independent production, but its location in the supply curve depends also price and inflow expectations, which affect the final generation scheduling. Javanainen (2005) concludes that different production types cannot be reliably estimated from supply curves which are based on aggregated production data published by TSOs.

Lastly, Vehviläinen and Pyykkönen (2005) present a model for the Nordic market, which combines aspects of fundamental and stochastic models. They argue that the fundamental variables are more stable in form and less complex to model than the spot price process itself. The spot price is represented as a deterministic function of the fundamental variables, which focus on the dynamics of the hydro-thermal market. Parameters are estimated from realized prices, consumption and production volumes as well as from historical climate data.

Chapter 4

An approach to short-term spot price forecasting

4.1 Proposed framework and key assumptions

4.1.1 Choice of approach

The literature review of electricity price models in the previous chapter indicated that approaches for short-term spot price forecasting fall into the categories of autoregressive (Section 3.4), artificial intelligence-based (Section 3.5) and fundamental models (Section 3.6). Autoregressive and artificial-intelligence based models are black-box techniques. Their main inputs are the past values of the variable to be predicted, which is the spot price. Demand or other market fundamentals may be also used as predictor variables.

Modelling the spot price directly with a black-box model would leave no room for applying knowledge about the structure of the electricity market. This prospect makes fundamental approaches attractive. An intuitive technique is to determine the price as the intersection of supply and demand functions, as proposed by Bunn (2000). Apart from short-term spot price models, this structure is present also in some stochastic models, such as Skantze et al. (2000), and Kanamura and Ōhashi (2007). A structural representation of supply and demand sides has several benefits. First, the representations of supply and demand can be broken further down into less complex sub-components, which can be modelled with most suitable techniques. Second, the analyst working with the model can adjust the supply and demand functions according to their own market view. Lastly, transparency of the model's workings is likely to improve its credibility and acceptance.

The main idea of the proposed framework is to construct a supply function and produce a price profile by projecting an exogenous demand forecast onto it, as inspired by Bunn (2000). He notes that in practical forecasting the supply function is assumed to be close to that of the previous day, and is adjusted based on any available special to reflect future supply situation. In this thesis, the supply function is estimated from data consisting of realized hourly system prices and production volumes. It represents actual pricing of supply on a certain day, including possible price distortions. The heuristic algorithm developed for the task is described in detail in Section 4.3.1. Due to the huge number of plants with different generation capacities and SRMCs, it would be extremely challenging to construct a system-level supply function from plant-level data. To begin with, part of the data is likely to be proprietary, and one would also have to estimate the water values of hydro power reservoirs.

One objective of this thesis was to study the relationship of market fundamentals and the spot price. The proposed framework attempts to explain shifts in estimated supply functions by changes in market fundamentals[1]. The estimated supply function represents only price-dependent production (PDP) sources, which in the Nordic case are reservoir hydro power and condensing production. If total production volumes were used in the estimation, the resulting supply function would be shifted by the temporal variation of baseload production, as illustrated in Figure 4.1. In this case, shifts of the supply function could not be attributed to changes in fundamentals alone. It is assumed that baseload production runs regardless of the spot price in the short run, and it is hence referred to as price-independent production (PIP). PIP does not satisfy the total demand alone, but some PDP is also needed. Thus, the spot price forecast can be calculated as the value of a PDP supply function at the point given by an exogenous PDP demand estimate.

The supply function is estimated from realized hourly price and PDP volume data. It is good to note that the realized PDP volumes reflect the market equilibrium, and hence the supplied PDP volume matches the demanded volume. The PDP demand estimate, which is used to determine the spot price in the forecasting period, is

[1]For an introduction to market fundamentals, refer to Section 2.6. Fundamental drivers of supply will be discussed thoroughly in Section 4.2.2.

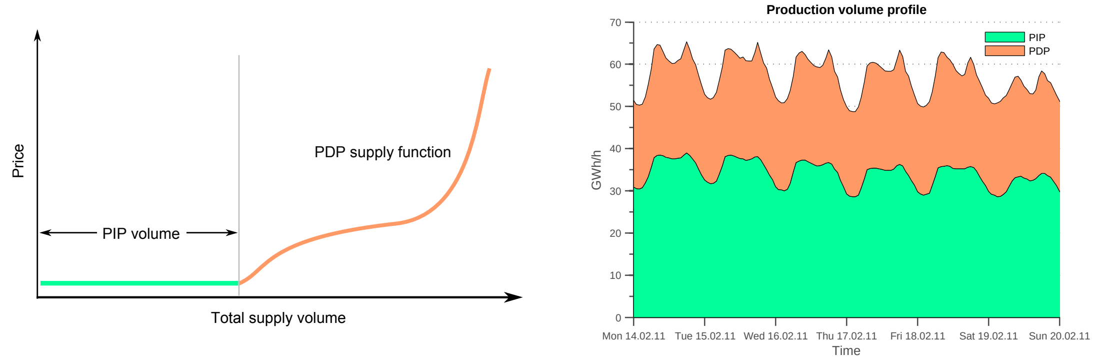


Figure 4.1: Illustration of how the profile of price-independent production (PIP) shifts the location of price-dependent production (PDP) in the supply function of total production, even though the pricing of PDP would not change.

derived from the power balance. It is assumed that the equation

$$Consumption = Production + Imports - Exports \tag{4.1}$$

holds for each hour. Here, $Consumption$ and $Production$ are total volumes of electricity consumed and produced within the Nordic market, and $Imports$ and $Exports$ are total volumes transmitted in/out of the Nordic market area. Furthermore, generation sources can be categorized as PIP or PDP. By setting $Production = PIP + PDP$, we can rewrite (4.1) to define

$$PDP := Consumption - PIP - Imports + Exports. \tag{4.2}$$

Hence, the estimated demand for PDP can be derived from estimates for total consumption, PIP, imports and exports. In this thesis, they are considered as exogenous variables. The actual sources for data will be discussed in Section 4.2.1.

The estimated supply function is assumed to be non-decreasing and piecewise-linear. Nord Pool Spot (2011c) requires that supply curves are non-decreasing, which is a reasonable requirement also for the estimated supply function. In this thesis, a single supply function is used for each hour of one day. In order that this type of supply function yields expected results, the pricing of PDP must be fairly unambiguous, meaning that same price is asked for same volume of PDP in each hour of the day. Otherwise the supply function cannot correctly represent pricing in each hour.

4.1.2 Key assumptions

In what follows, assumptions implied by the proposed framework are summarized.

The system price is determined by matching aggregated supply and demand curves for each hour. As described in Section 2.4.1, the Nordic system price is determined by an implicit auction. The assumption somewhat simplifies the actual pricing mechanism. In addition to hourly sell and purchase orders, participants can also specify flexible hour offers and block orders. The hourly orders are matched by intersecting aggregated supply and demand curves for each hour, which yields an equilibrium price and volume. Flexible hour offers and block orders are activated if they improve the social welfare given by the basic allocation. When they are activated, the affected hours have a new equilibrium price. In case of a block order, prices in the hours covered by the block depend on each other, which is contrary to the assumption.

Physical consumption, production and exchange volumes act as a proxy for Elspot supply, demand and flow quantities. The system price is determined by the prices and volumes specified in bids submitted to Elspot, but the proposed framework uses physical consumption, production and exchange volumes to approximate these financial quantities. This is necessary, as the framework depends on distinguishing the contribution of different generation sources in total production. Elspot volumes are closely related to physical volumes, because spot trades involve physical delivery of power. In 2011, Elspot's daily turnover amounted on average to 76 % of the physical turnover. Moreover, correlation between Elspot and physical turnover was estimated to be 0.99, which strongly supports the validity of the assumption.

Demand is inelastic on short term. Inelasticity of demand is commonly assumed in spot price models, such as in fundamental approaches proposed by Skantze et al. (2000) and Vehviläinen and Pyykkönen (2005), and in autoregressive models described by Nogales et al. (2002) and Weron and Misiorek (2008). Furthermore, Bye and Hansen (2008) studied elasticities in Norway and Sweden and concluded zero elasticity in the summer and very low elasticities in the winter.

Generation can be divided to price-dependent (PDP) or price-independent production (PIP). PDP plants have considerable short-run marginal costs (SRMC), and in the case thermal generation, also start-up costs. A PDP plant runs only when the market price covers these costs. On the other hand, the SRMC of a PIP plant is remarkably lower than the normal price level, and the production decision is independent of the market price. As described in Section 2.2, the major generation sources in the Nordic countries are hydropower, nuclear power, wind power and conventional thermal generation. Hydropower can be further divided to regulated and unregulated generation. In conventional thermal generation, a distinction can be made between CHP and condensing plants. Out of these sources, condensing and regulated hydropower are considered as PDP and the rest as PIP.

The market is perfectly competitive. Under perfect competition, the market price is equal to the marginal cost of the last unit produced. If the assumption holds, the price should reflect fuel and emissions costs, or water value. As discussed in Section 2.5, the marginal cost of hydro power is the water value. Therefore, the price should be connected to the market fundamentals which affect the water value. A company exploits market power if it attempts to manipulate prices in order to raise its profits. A number of studies have assessed the presence of market power in the Nordic market. Kauppi (2009) argues that a part of Nordic hydro power resources are operated strategically, but under most circumstances the concentration of market power is not high enough to affect the price level. In their review of empirical market power studies, Fridolfsson and Tangerås (2009) find no evidence of systematic exploitation of the Nordic market power on the system level.

4.1.3 Specification of the framework

The proposed framework is essentially summarized by the following steps, which describe how the system price forecast is produced. The steps are outlined in Figure 4.2.

1. Estimation data set is built from hourly realized prices and PDP volumes.
2. PDP supply function is estimated from the data set using a heuristic algorithm (described in Section 4.3.1).
3. For each day for which a forecast is to be made, the supply function estimate is adjusted based on any information on market fundamentals.
4. Demand for PDP is estimated for each day for which a forecast is to be made. The PDP demand estimate is considered as an exogenous variable.
5. System price forecast is calculated as the values of the adjusted supply functions at the estimated PDP demand.

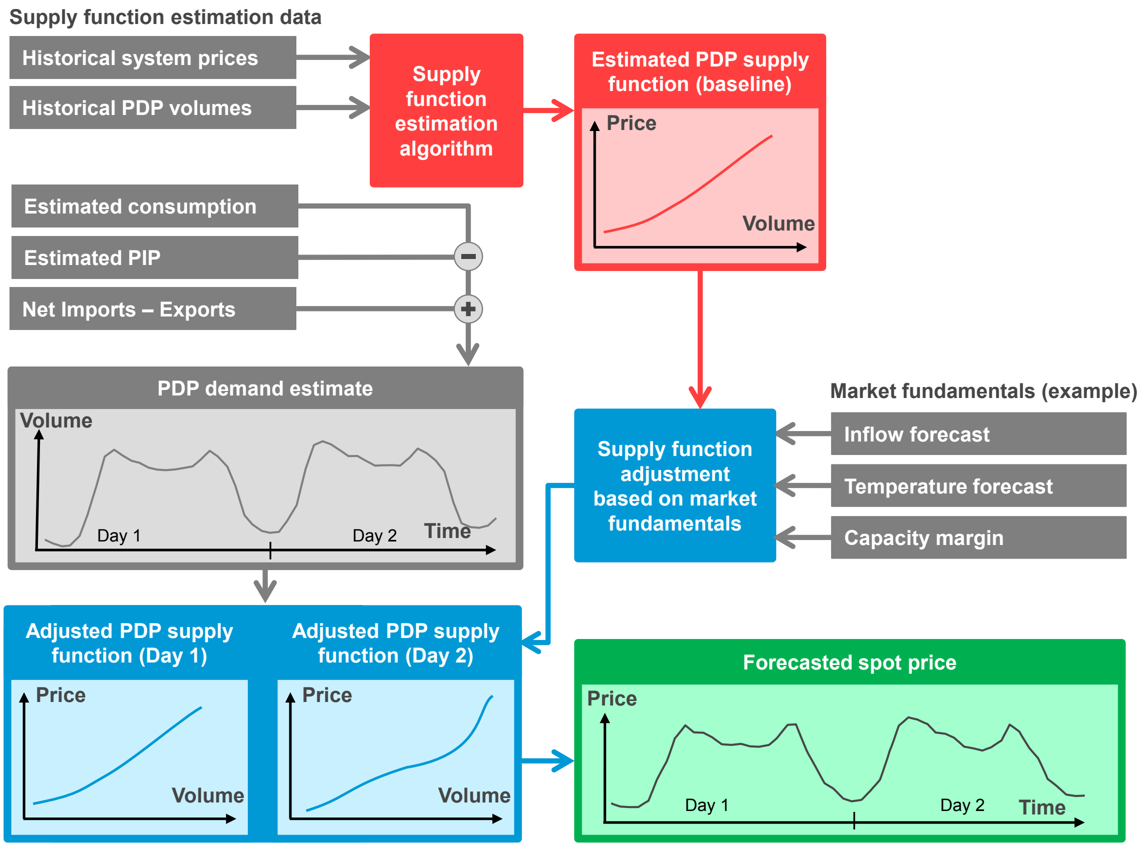


Figure 4.2: Outline of the proposed forecasting framework. Exogenous input variables are in grey colour. The data is will be discussed in Section 4.2. The supply function estimation algorithm (red) is described in Section 4.3.1. Changes in the estimated supply function with respect to market fundaments (blue) will be discussed in Sections 4.3.2 and 4.3.3.

The framework relies on several data sources to determine the volume of PDP. The realized PDP volume can be directly calculated by summing up price-dependent production volumes from appropriate statistics. However, it is significantly harder to create a direct forecast of the PDP demand. Hence, the forecasted PDP demand is derived from the power balance equation as in (4.2). Consumption and PIP are independent of price, and can therefore be modelled separately or acquired from an external source. Exchange between the Nordic market and Central Western Europe is determined in a so-called market coupling process, which sets socially optimal power flows on the interconnections (European Market Coupling Company, 2009). Connections to Russia have no coupling mechanism, but the exchange depends solely on the difference between Russian price and Finnish area price. Nevertheless, also estimates for exchange can be obtained from external sources.

4.2 Description of data

The forecasting approach developed in this chapter is somewhat data-intensive. Exhaustive production data per generation source are required to determine the volume of price-dependent production (PDP). Furthermore, market fundamentals are used to adjust the estimated supply function. The first part of this section describes actual data sources for the Nordic market, and the second part discusses market fundamentals and their causal effects on the pricing.

4.2.1 Data set

The proposed framework relies on three main categories of data, which are system price, production volumes and market fundamentals. Realized system price and PDP volume are used to construct a supply function estimate, and the price forecast is given as the value of the estimated supply function at the forecasted PDP demand. The estimated supply function is adjusted according to relevant changes in the market fundamentals.

The data can and needs to be obtained both from public and commercial sources. The system price is published by Nord Pool Spot. Actual and historical production data per hour and generation source are published by Nordic TSOs, but it may occur with a lag that essentially decreases the usability of the data. Fingrid (FI) and Statnett (NO) publish production statistics in real-time, but in the case of Svenska Kraftnät (SE) and Energinet.dk (DK) the lag ranges from one week to several months. Therefore, it is necessary to utilize the services of commercial data providers. Commercial data is also needed to forecast the demand of PDP, unless the forecaster has capability to model it themselves. Finally, weather and hydro-related fundamentals are typically bought as a service due to highly complex models behind them.

The framework is based on the volume of PDP instead of total production, as explained in Section 4.1.1. Wind, unregulated hydro power, nuclear power and CHP are considered as PIP, and condensing and regulated hydro power as PDP. The hydro power production volumes provided by TSOs do not differentiate unregulated and regulated hydro production. For that reason, a crude generalization has to be made so that Finnish and Swedish hydro power is considered as PIP, and Norwegian hydro power as PDP. This division is supported by Javanainen (2005), who studied publicly available, aggregated production data, and found that Norwegian hydro production showed stronger price-dependency than Swedish and Finnish hydro production. The division between PIP and PDP should definitely be improved at a later stage. The division used in the empirical part of this thesis is presented in Table 4.1. Estonia was not included in the data set used in this thesis. It has practically no effect on the analysis which is conducted on system level, as Estonia makes up only 2–3 % of total Nordic production.

Table 4.1: Division of generation sources (as reported by the TSOs) into price-independent (PIP) and price-dependent production (PDP). Because regulated and unregulated hydro power were not differentiated, the classification of hydro was based on the relatively higher degree of regulation in Norway compared to Sweden and Finland.

Country	PIP	PDP
Finland	Hydro, Nuclear, CHP (industrial), CHP (district heating)	Condensing
Sweden	Wind, Hydro, Nuclear, CHP	Condensing
Denmark	Wind, Local CHPs	Central power stations*
Norway	—	Hydro

*Includes both district heating CHP and condensing power plants.

4.2.2 Fundamental drivers of supply

An estimated supply function represents pricing in the market at certain point of time. It is assumed that the pricing depends on the prevailing market fundamentals (Section 2.6). Forecasting requires that the supply function is adjusted in proportion to changes in the fundamentals. This subsection presents main fundamentals and discusses their causal effects on electricity pricing. Market data from years 2009–2011 is used to illustrate the effects, which are summarized in Table 4.2. The argumentation builds strongly on the concept of *water value* and the dynamics of hydro and thermal generation, which were discussed in Section 2.5.

Volume of price-dependent production (PDP) has a strong effect on the price level and the periodic variation of the system price. PDP plants have different marginal production costs depending on the technology and other plant-specific factors. When the demand for PDP increases, generation capacity with a higher marginal production cost needs to be activated. There can be big differences between the marginal costs of peaking power plants, which adds to the price volatility during highest demand. Under the competitive market assumption, the plant with the highest marginal cost sets the market price. Furthermore, PDP reflects the periodical variation of consumption, which takes place on intra-day, weekly and yearly levels. A measure of consumption or demand is included in the price formula of fundamental models (Section 3.6) or widely used as a predictor variable in time-series models (Section 3.4).

Short-run marginal cost (SRMC) of condensing production has a critical role in the allocation of hydro and condensing production. As defined in Section 2.2.1, the main factors of SRMC are fuel cost, emission cost and plant efficiency. It should be noted that start-up costs must be also taken into account. The SRMC of an average coal condensing plant is considered for price forecasting purposes. Marginal costs are taken into account in long-term models, such as EMPS (Wolfgang et al., 2009), which are used for determining water values. When water values are below SRMC of coal condensing, hydro power will precede condensing production in the merit order. Subsequently, the spot price will not exceed the SRMC of condensing production if the amount of hydro power offered to the market is enough to satisfy the demand without condensing power.

Temperature has a non-linear effect on consumption. When below the level where heating is needed, power consumed by electrical heating is proportional to the temperature. Likewise, but very infrequently in the Nordic countries, electrical cooling devices begin to increase consumption above a certain temperature level. Especially in spring, rising temperatures speed up the melting of snow and ice, which contributes to inflow.

Inflow measures the power generation potential of water, and is typically given in energy units. Fundamental Nordic price models commonly use inflow values aggregated per price area or the whole market. According to Laukkanen (2004), inflow models take temperature, precipitation, snow reservoirs and soil water as input. Javanainen (2005) studied hydro production in Norway and concluded that inflows are likely to result in uncontrolled (price-independent) hydro production in areas where the production system has low flexibility. As Norway has generally most flexible hydro production capability in the Nordic market, it can be assumed that hydro production in Sweden and Finland depends even more strongly on inflows. As a result, inflows increase PIP, which can be expected to lower the demand for PDP. In the face of spillage, inflows can increase the production pressure on price-dependent hydro and thus lower the pricing.

Hydro balance represents the deviation of current water resources from a multi-year average at the same time of the year. The balance calculation includes water in hydro reservoirs, snow and soil. Vehviläinen and Pyykkönen (2005) regard hydro balance as an indicator of production willingness in their fundamental model. A surplus of water resources can limit producers' flexibility and lower the water values. On the other hand, scant water resources may allow a highly flexible production planning, which can lead to high prices and profits from hydro generation. The Nordic hydro balance is the sum of regional hydro balances, which can differ depending on the local conditions.

As can be seen in Figure 4.3, peaking system prices have occurred during high PDP demand in the first quarters of 2010 and 2011. During the second and third quarters PDP demand was typically at minimum level, and the system price was somewhat following the level of SRMC Coal. The price profile in these quarters is considerably flatter than that of PDP. It is likely that price-dependent hydro producers are able to keep their water values close to, or even above SRMC Coal. Moreover, it can be observed that most price collapses coincide with sudden increases in inflow. During Q3 in 2011 prices collapse below 10 EUR/MWh level, which is possibly due to the combination of high hydro balance and a rise in inflow. Prices seem to generally recover faster from sudden downward movements than upward moments.

Tightness, or the ratio of actual production to available production capacity of Norwegian hydro production is examined as another indicator of spiking prices. Javanainen (2005) argues that the aggregated supply curve of Norwegian production is increasing due to hydro power plants having different water values. Differences between plant-specific water values can be substantial towards the high end of the supply curve. Therefore, high tightness can be a signal for increased price volatility. It can be seen in Figure 4.4 that tightness follows the profile of PDP, but has a larger relative difference between working day and weekend levels. Moreover, extremely high price volatilities appear, when tightness is above 70 %.

To conclude, historical values were used to analyse the relationship of key fundamentals and system price. The use of historical values is typical for long-term fundamental analyses of the power market, such as Johnsen (2001) and Botterud et al. (2010). However, hydro production planning relies much on the outlook of

Table 4.2: Summary of fundamentals affecting the pricing of PDP supply on system level in the Nordic market

Fundamental	Effect on pricing	Data frequency	Main arguments
Hydro balance	↘	week	Indicates the level of control that hydro producers have over regulated (reservoir) production
Inflow	↘	day	Increases unregulated hydro production, and may increase production pressure of regulated hydro
Precipitation	↘	day	Input of inflow calculation
Temperature	↘	day	Inverse relation to consumption may affect water valuation. Input of inflow calculation
SRMC Coal	↗	day	Direct effect on the pricing of condensing production. Reflects the opportunity cost of withdrawn hydro production
Tightness	↗	hour	The higher the utilization, the more expensive plants are brought on-line

weather fundamentals. Uncertainty in weather forecasts may lead to ex-post suboptimal water values and production allocation. Example 1: A producer allocates water to be used in the following weeks when low precipitation is forecasted. If the forecast later turns wet, the producer may earn less profit due to increased run-of-river production. Example 2: A producer allocates water for week 1 in January, because temperatures are expected rise in week 2. Later the temperature forecast for week 2 turns considerably colder, which increases the consumption forecast. Consequently, water values for week 1 were set too low.

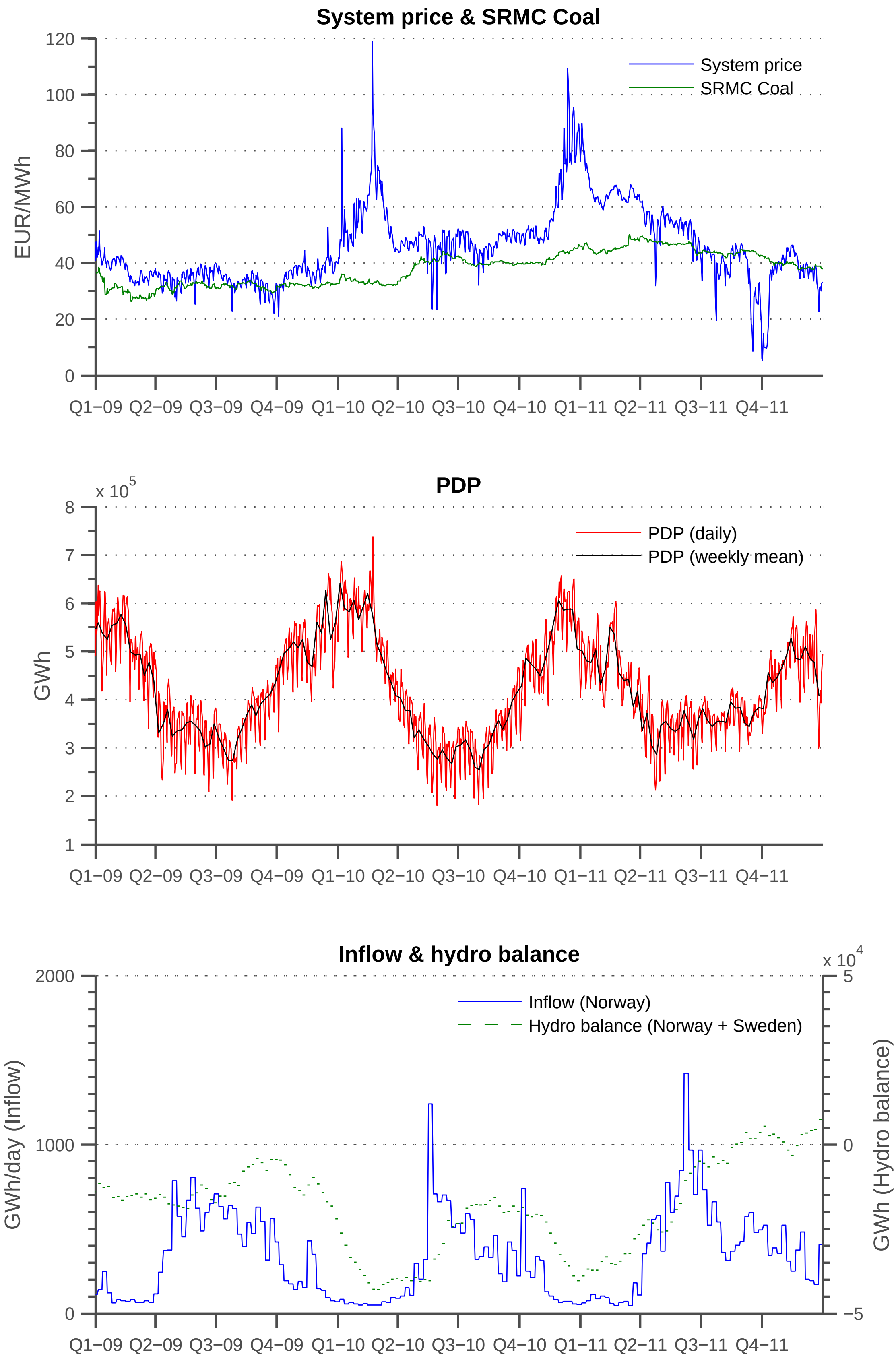
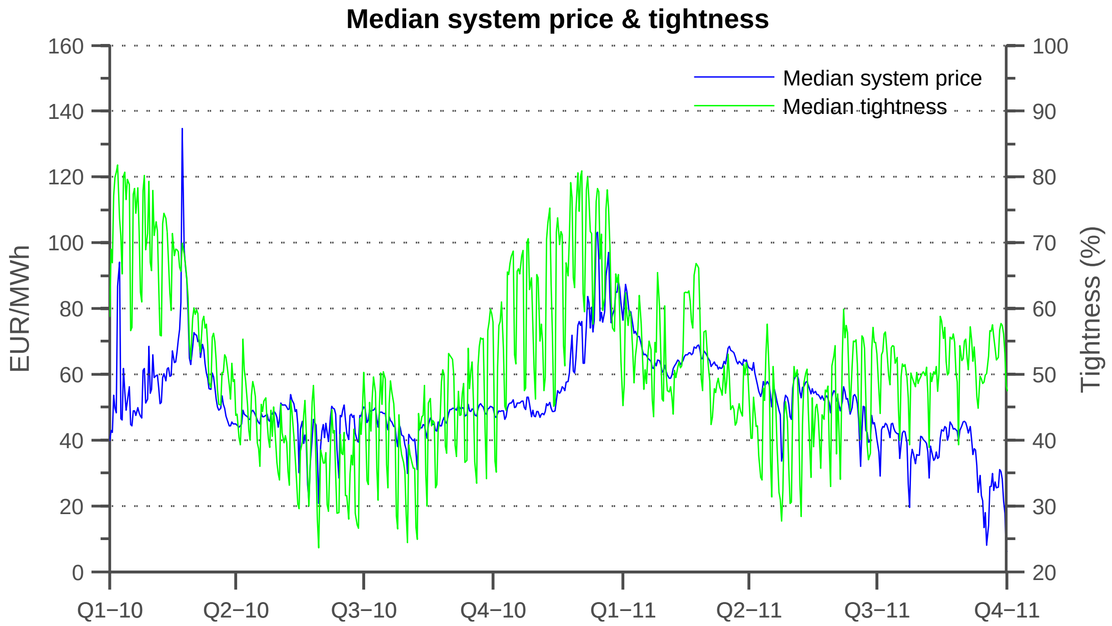


Figure 4.3: Historical values of the median system price, SRMC Coal, inflow, hydro balance and PDP



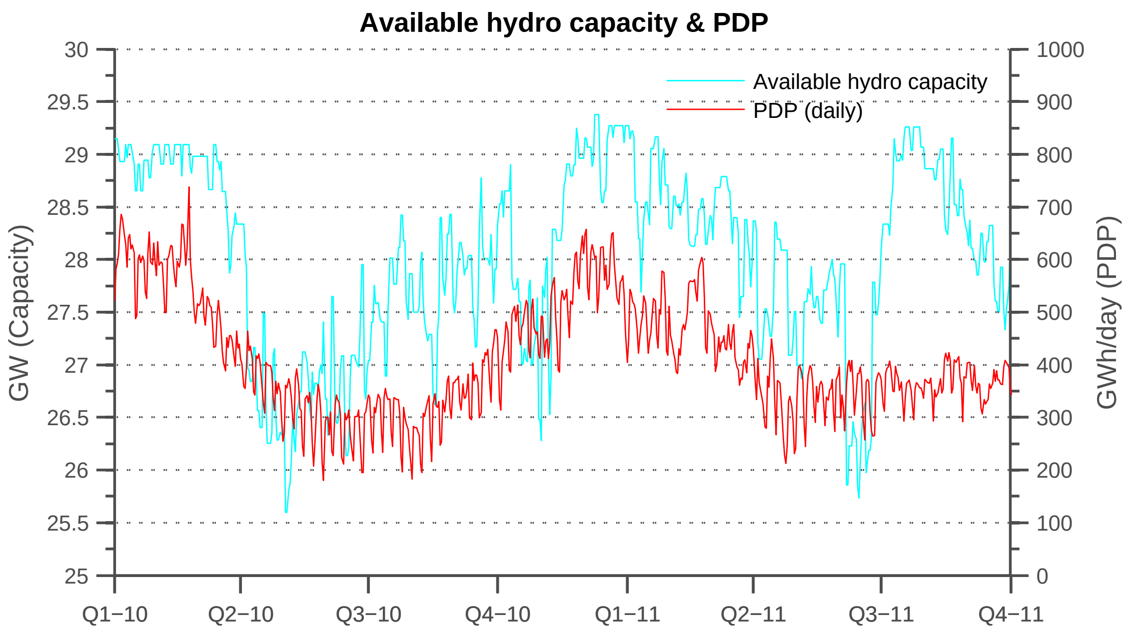


Figure 4.4: System price and utilization of Norwegian hydro power capacity (*Tightness*)

4.3 Validation of the framework

This section describes how the proposed forecasting framework was validated against historical market data with a backtesting procedure. The data set used for validation is described in Section 4.2 and spans over the years 2009–2011. Because PDP volume is an external variable in the framework, the validation focuses on the estimation of the supply function and the effects of changes in market fundamentals. The validation procedure used realized PDP volume both for estimating the supply function and for calculating the forecast price. In effect, the forecast error, defined as the difference of realized and forecasted price, was completely attributed to the inaccuracy of the supply function.

4.3.1 Estimating the supply function

The PDP supply function represents the pricing of electricity in the Nordic market on a certain day. As part of this thesis, a heuristic algorithm was developed to estimate a supply function from data consisting of observed hourly system prices and corresponding PDP volumes. The algorithm approximates the data with a first degree spline. It is basically a piecewise defined function that connects the data points chosen as *knots* with line segments. Furthermore, it is assumed that the supply function is non-decreasing.

Constructing a supply function is unambiguous if each PDP volume in the data set is associated with a certain price at all times. In this case, the plotted data points align to a single curve as illustrated by the upper curve in Figure 4.5. On the other hand, it is possible that a certain volume has been traded at different prices during the same day, as illustrated by the lower curve in Figure 4.5. In that case the supply function cannot represent correct pricing for each hour of the day. The estimation algorithm deals with the unambiguity by prioritizing more recent observations over less recent ones.

The estimation algorithm has three major phases. In Phase I, the data is split into fixed-size volume intervals, and the most recent observation from each interval is stored. In Phase II, stored observations belonging to the most recent calendar day are selected as knots. In Phase III, older stored observations are selected as knots as long as their inclusion does not violate the requirement for a non-decreasing function shape.

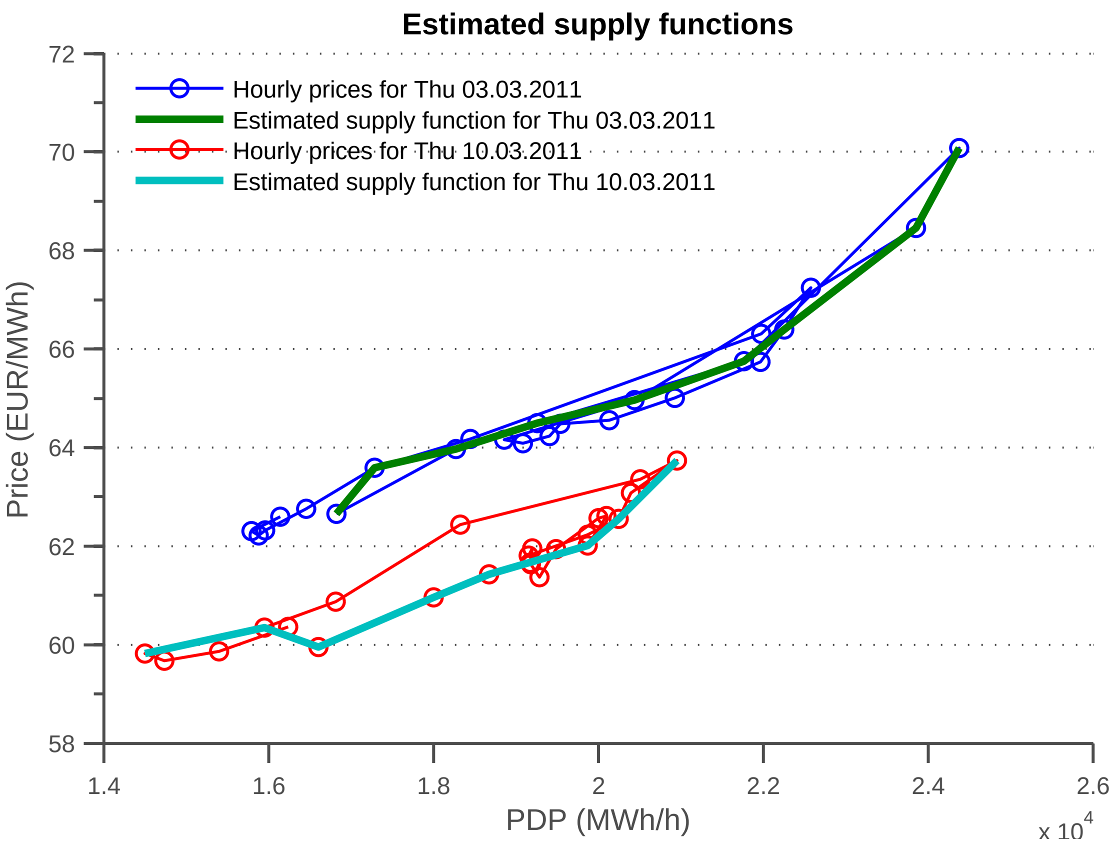


Figure 4.5: Illustration of unambiguous and ambiguous pricing. On 3.3.2011 same volumes have been traded at fairly similar prices in all hours. In contrast, on 10.3.2011 there is a 1.20 EUR/MWh difference in the volume range of 18 GWh/h between different hours.

Let $x_i = \{t_i, q_i, p_i\}$ denote an observation at time (hourly resolution) t_i, PDP volume q_i and system price p_i. If $t_i < t_j$, then time t_i precedes t_j. Given an estimation data set $X = \{x_i\}$ and volume interval size v_{int}, the algorithm is defined as follows:

PHASE I

Set $X_{MostRecent} = \emptyset$

Set lower volume bound $l = \min\{q_i\}$

repeat

Set upper volume bound $u = l + v_{int}$

Add x_k to $X_{MostRecent}$ s.t. $t_k = \max\{t_i \,|\, l \leq q_i \leq u\}$

Set $l = u$

until $l > \max\{q_i\}$

PHASE II

Define $D :=$ “set of calendar days for which there are observations in $X_{MostRecent}$”

Set $d_{max} = \max D$

Define $X^{d}_{MostRecent}$:= "observations in $X_{MostRecent}$ belonging to calendar day $d \in D$"

Set $X_{SupplyFunction} = X^{d_{max}}_{MostRecent}$

PHASE III

for each $d \in D \setminus \{d_{max}\}$ **do**

Iterate d from the most recent to the least recent day

for each $x_i \in X^{d}_{MostRecent}$ **do**

Iterate x_i from the smallest to the largest volume q_i

if $q_i < q_j$ & $p_i < p_j$ $\forall$ $x_j \in X_{SupplyFunction}$

or $q_i > q_j$ & $p_i > p_j$ $\forall$ $x_j \in X_{SupplyFunction}$

or $\exists$ $x_j \in X_{SupplyFunction}$

s.t. $q_j < q_i$ & $p_j < p_i$ & $q_i < q_{j+1}$ & $p_i < p_{j+1}$

then

Add x_i to $X_{SupplyFunction}$

end if

end for

end for

Return $X_{SupplyFunction}$

It should be emphasized that the estimation algorithm was developed primarily for the purpose of evaluating and validating the proposed framework against historical market data. It enabled automatic and non-interactive estimation of supply functions, which was necessary as several years of historical data was processed. In a live forecasting situation more time can typically be invested into the estimation process, and the quality of the estimate could be improved by checking the unambiguity of pricing and by making manual adjustments to the estimated function.

4.3.2 Forecasting with a static supply function

As the first stage of validation, prices were predicted using a single PDP supply function for the whole forecasting period. A situation was considered, where the supply function is estimated from most recent realized price and PDP volume data using the algorithm presented in the previous subsection. The so-obtained supply function was used to statically represent pricing of PDP throughout the forecasting period. The forecast price was determined as the value of the supply function at the demand of the moment.

A forecast is typically done before Elspot's gate closure at 12:30 CET, so that the results can be used to adjust the bids. Spot prices for the current day are known at that time, and theoretically it is possible to observe realized PDP volumes up to the last full hour. In reality, the availability of PDP data varies depending on the data providers, as noted in Section 4.2.1. If it is desired to use only realized data for supply function estimation, a lead time of one day should be assumed between the estimation and forecasting periods. The lead time can be reduced by extending the estimation data set with actual prices and PDP volumes estimated for the non-realized hours of the current day.

In what follows, observations from the validation experiments will be discussed. The level of PDP in validation data was highly volatile, and it was common that PDP demand in the forecasting period could go out of the range where the PDP supply function was defined. This was usually connected to changes in total consumption. One reason were differences in consumption levels during working days and the weekend, and another reason were changes in outside temperature, which affected the load from electrical heating. Possible solutions for the issue are extrapolating the supply function, or extending the estimation period so that it would cover the desired volume range. Both solutions have their limitations.

Extrapolation of the supply function, especially to higher volumes, depends essentially on the maximum available capacity of PDP. It is likely that best extrapolation can be achieved by manually extending the supply curve, while taking fundamentals into account. Given that the additional supply data is consistent with the more recent data, extending the estimation period extends the volume bounds, where the supply function is defined. As pointed out in Section 4.3.1, the estimation algorithm first uses most recent data, and then adds data points which are consistent with the increasing slope of the function. In this validation procedure, the estimation period was extended to match the volume range in forecasting period. If PDP values were still out of bounds, no price was forecasted for the concerned hour.

The accuracy of forecasts with a static supply function degraded rapidly as the lead time increased. The effect was studied by forecasting the price one day ahead with lead times of 0, 1 and 3 days. The length of the estimation period was 21 days, but the estimation algorithm typically exploited only 2–5 most recent days. A number of time periods were selected from the data set to represent different seasons. As illustrated by Figure 4.6, short lead times generally produced more accurate forecasts. Because the validation was carried out by using realized PDP values both for supply function estimation and forecasting, there is no uncertainty in the demand side. Therefore, all forecast errors were viewed as an indication that the static supply function did not represent true pricing at the time of the forecast. If the pricing was unambiguous, the forecast error could be reduced by adjusting the supply function. On the other hand, the forecast error can also be due to ambiguous pricing, which cannot be represented by the supply function.

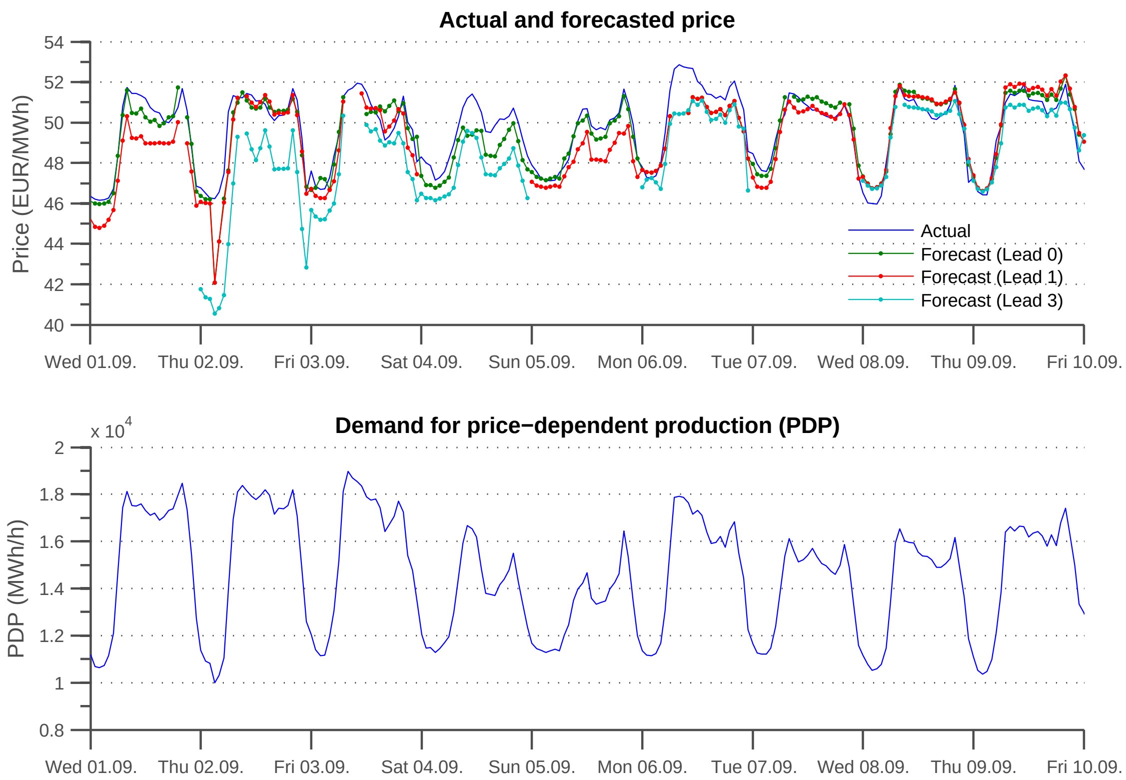


Figure 4.6: The upper plot shows the realized price and forecasts with different lead times. The lower plot shows the corresponding demand for PDP. The data is from September 2010.

The unambiguity of pricing was checked as part of validation. Even if the piecewise-linear supply function is incapable of representing the actual pricing, *price trajectories* of realized hourly prices and PDP volumes can be used to study shifts in the pricing of supply. Figure 4.7 illustrates the approach with the same time period as in Figure 4.6. It can be seen that pricing was relatively unambiguous from 30.8. to 1.9. except for the highest volumes. The clearest example is 30.8., when the trajectory formed a small 'loop' in the high end of the volume range: The volume was about 17 GWh during two hours, but in one hour the price was around 49 EUR/MWh, whereas it was about 1.5 EUR/MWh more during the other hour. It can be also seen that the supply curve shifted to left from 30.8. to 31.8., which increased prices roughly by 1.5 EUR/MWh. The supply curves for 2.9. are heavily based on the most recent realized data. In Figure 4.6, price forecasts with lead times 0 and 1 match the actual price profile in the middle of the day and evening hours, but underestimate the price in morning hours. This can be explained by the realized price trajectory of 2.9., where the upper part of the ‘loop’ corresponds to the morning hours, but the supply curve estimates are closer to the evening hours.

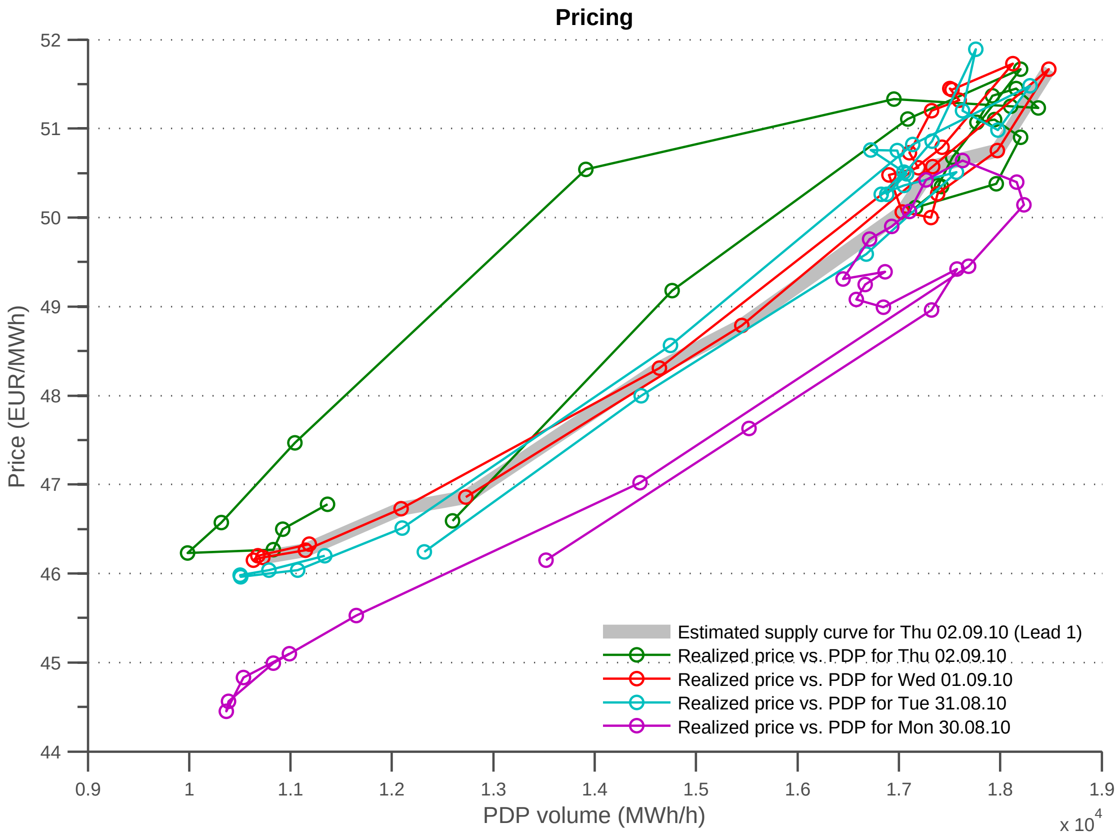


Figure 4.7: Realized price trajectories and the estimated supply function (lead 1) from the forecasting period shown in Figure 4.6

To conclude, backtesting with a static supply function indicated that the pricing of PDP is highly dynamic. The idiosyncratic intra-day price profile with distinct price levels for night and day hours originates from the profile of PDP demand. The supply function practically translates demand levels into price levels. Dynamical pricing supports the use of a relatively short period of data for estimation and suggests that the price forecast should rely on the estimated PDP volume. Some implications can be drawn for other types of models. A time series ARIMA model with PDP as exogenous variable may yield satisfying results when the level of PDP volume remains stable. However, the model is linear with respect to PDP and may not capture steeply increasing prices, which occur when PDP approaches the limit of available production capacity. This problem could be alleviated by applying a more sophisticated ARIMA-GARCH model, which could measure the dynamic volatility (Section 3.4). An artificial neural network model could possibly benefit from a long estimation data set, which can be used to train the model (Section 3.5). It may also be able handle the intra-day dynamics better than a static supply function. However, the proposed framework provides higher transparency over price formation. Finally, it should be emphasized that the framework depends heavily on the quality of the estimated supply function.

4.3.3 Analysing the effects of changing fundamentals

As the second stage of validation, the forecast error of the static PDP supply function was studied with regression analysis. Analysis of supply functions showed that the pricing of PDP varies in relation to time. Under the assumption of competitive pricing, the variations can be connected to fundamental price drivers discussed in Section 4.2.2. The regression models studied in this subsection attempt to explain the realized price by correcting the output of the static supply function with fundamental predictors. The estimation algorithm produces supply functions, which resemble pricing at the end of the estimation period. Parameters were estimated by fitting the regression model to the whole estimation period using ordinary least squares.

The proposed framework operates on hourly frequency, and the fundamental predictors should likewise have a high frequency to distinguish hours. fundamentals with weekly frequency, such as actual hydro balance, are effectively constant within the forecasting horizon. Out of the fundamentals discussed in Section 4.2.2, the following were included in the regression analysis: inflow in Norway, consumption-weighted average temperature in Norway and Sweden, supply tightness in Norway, and the system-level PDP. Inflow and temperature had daily frequency, while tightness and PDP were specified on hourly level.

Experiments were made with three different regression models. Linear models were chosen because they can approximate non-linear effects, and parameter estimation is simple. A fundamental predictor for day d and hour h was defined as

$$x_{d,h} = f_{d,h} - \bar{f}, \tag{4.3}$$

where $f_{d,h}$ is the actual value of the fundamental, and $\bar{f}$ is a base value from the estimation period. $\bar{f}^{Inflow}$ and $\bar{f}^{Temperature}$ were defined as the average of days, from which data points were included in the supply function estimate (typically 2–3 days). In order not to lose a level of detail, the base value for tightness was defined as $\bar{f}^{Tightness} = \bar{f}_h^{Tightness}$, the average value of hour h over days, from which data points were included in the supply function estimate.

In the first experiment, pure differences were used as predictors, resulting in the model

$$p_{d,h} = \beta_0 + \beta_1 p_{d,h}^0 + \beta_2 x_{d,h}^{Inflow} + \beta_3 x_{d,h}^{Temperature} + \beta_3 x_{d,h}^{Tightness} + \varepsilon_{d,h}, \tag{4.4}$$

where $p_{d,h}^0$ is the price forecast given by the supply function estimate at PDP demand $q_{d,h}$, β_i are the regression coefficients and $\varepsilon_{d,h}$ is the error term. The second experiment used interactions of PDP and the fundamental predictors as variables, specified as

$$p_{d,h} = \beta_0 + \beta_1 p_{d,h}^0 + q_{d,h} \left(\beta_2 x_{d,h}^{Inflow} + \beta_3 x_{d,h}^{Temperature} + \beta_4 x_{d,h}^{Tightness} \right) + \varepsilon_{d,h}. \tag{4.5}$$

The third experiment extended the interaction model by adding future values of fundamental predictors. Future values of 1–5 days ahead were used with weights identical to the actual value. With predictors up to 5 days ahead, the model was

$$\begin{aligned} p_{d,h} = \beta_0 + \beta_1 p_{d,h}^0 + q_{d,h} \Big(& \beta_2 x_{d+0,h}^{Inflow} + \cdots + \beta_7 x_{d+5,h}^{Inflow} \\ & + \beta_8 x_{d+0,h}^{Temperature} + \cdots + \beta_{13} x_{d+5,h}^{Temperature} \\ & + \beta_{14} x_{d+0,h}^{Tightness} + \cdots + \beta_{19} x_{d+5,h}^{Tightness} \Big) + \varepsilon_{d,h}. \end{aligned} \tag{4.6}$$

The regression models were backtested against the validation data set. Time periods were selected to represent different seasons and fundamental situations, such as periods when fundamentals are very stable or change rapidly. Length of the estimation period was 14 days, and the forecast was made 7 days ahead with no lead between estimation and forecasting periods. For temperature and tightness, realized values were used both in the estimation and forecasting periods. For inflow realized values were available only on weekly level. Hence, values from a time series of latest forecasts for each day were used instead.

Conclusions were made by analysing the error terms and the estimated regression coefficients. It was checked if the predictors were statistically significant and if their signs corresponded to the expected fundamental effects, which are summarized in Table 4.2. Overall, the interaction model (4.5) performed better than the basic model (4.4). This could be expected, as correlation between the realized price and PDP demand was noticed in the analysis of supply curve behaviour (Section 4.2.2). The third model (4.6) produced a high goodness of fit due to the high number of variables in the model, but it did not produce better forecasts then the interaction model.

The regression analysis did not produce reliable quantitative results, which could be used to adjust the price forecasted by an estimated supply function. However, the following conclusions could be made based on the analysis. The regression coefficients could not be regarded as constant during the estimation and forecasting periods, which together spanned 21 days. Best results were produced under conditions, where the values of the fundamental predictors in the forecasting period remained within the same limits as in the estimation period. If the values changed significantly, the model typically under or overestimated the effect. This was particularly noticed with inflow and temperature. In the studied models, the effect of a fundament was directly proportional to its value difference. However, the findings suggest that the effects are non-linear even on the very short term. In several cases the price given by the regression model was inferior to the output of the static supply function, because fundamental predictors adjusted the price in the wrong direction.

4.4 Facilitating price forecasts in practice

This section sums up experiments conducted with the proposed forecasting framework and puts forward how the framework can benefit day-to-day forecasting activities.

Realized volumes of price-dependent production (PDP) are needed for supply function estimation. They can be calculated from the power balance as in (4.2), or by directly summing up PDP volumes from appropriate statistics. Hourly production statistics per generation source are required in both cases, and in principle the data is publicly available. Nord Pool Spot publishes total consumption, production and exchange volumes, and each TSO reports hourly production statistics per generation source. However, the Swedish and Danish TSOs publish production statistics with a considerable lag, which can be several months in the former case and about 10 days in the latter. In effect, the public data is of little value, and the forecaster must turn to commercial data providers, which are capable of supplying up-to-date figures.

The estimate of future PDP demand is an external variable in the framework. Similarly to the realized values, the estimate can be derived from the power balance by using forecasted values, or an external PDP forecast may be used. The forecaster should take into account that the estimate for PDP implies certain expectations on consumption, exchange and PIP. Moreover, the exchange implies a certain price difference between markets. Major adjustments in the supply function may affect the price difference to such extent that the direction of the exchange reverts, which could in turn change the whole PDP estimate. Furthermore, the estimate should reflect reductions in consumption, production or available transmission capacity events as indicated by Nord Pool Spot's urgent market messages (UMMs).

A PDP supply function that represents the actual pricing of a particular day is the basis for building estimates for following days. The estimation algorithm presented in Section 4.3.1 can be used to produce a crude estimate, which should be reviewed and tuned by the forecaster. In particular, it should be checked whether the pricing has been unambiguous. If not, special consideration is needed, because a single supply function cannot represent the price in each hour of the day. Moreover, the forecaster must extrapolate the supply function estimate to cover the volume range of PDP demand in the forecasting period, if the future demand differs from the volumes observed in the estimation period.

Statistical modelling of short-term fundamental effects turned out to be extremely challenging. Supply curves can be used to measure changes in pricing, but it is up to the forecaster to connect the changes with market fundamentals. Likewise, the forecaster should take a view on the fundamentals and adjust the supply function estimates for each day in the forecasting period.

Chapter 5

Conclusions

The purpose of this thesis was to develop and evaluate a framework for forecasting hourly spot prices of electricity in the Nordic market up to one week ahead. This type of short-term price forecasts can be used to support operational planning and electricity trading decisions. Only a few references related to fundamental short-term price forecasts in the Nordic market were found (Dueholm and Ravn, 2004; Javanainen, 2005). Overviews of electricity price forecasting can be found e.g. in Bunn (2000), Bunn and Karakatsani (2003) and Weron and Misiorek (2006).

The research method of this thesis consisted of two parts. First, Nordic market data was studied in order to identify fundamental drivers affecting price formation, and second, a forecasting framework was designed. The empirical part of the thesis consisted of implementing the framework and validating it against actual market data from years 2009–2011. As a major part of the implementation, an algorithm was created to estimate aggregated supply functions from market data. Finally, an effort was made to explain changes in estimated supply functions by changes in market fundamentals.

The framework proposed in this thesis was based on a fundamental approach inspired by Bunn (2000), who argues that a short-term price profile can be forecasted by constructing a supply function and projecting a demand estimate onto it. In this this thesis the approach was developed further to match the idiosyncrasies of the Nordic market. The supply function was estimated from realized prices and production volumes. Thus, it represented actual pricing on a certain day, including possible price distortions due to market inefficiency or other reasons. The production and consumption volumes considered only price-dependent production (PDP). If total volume of production would be used, the estimated supply functions would be shifted by temporal variation in the volume of price-independent production (PIP). When only PDP volumes were used, shifts in the supply functions could be attributed to changes in water values, short-run marginal costs and related market fundamentals.

The empirical results showed that the spot price was highly volatile and that the effects of market fundamentals have complex dynamics. These effects arise from the supply side, whereas the short-term demand is quite predictable. The price trajectories of realized PDP volumes changed from day to day, and occasionally the pricing differed between the hours of a single day. The supply function estimation algorithm exploited typically data from the previous 2–5 days, while older prices were incompatible with the requirement for a non-decreasing supply function estimate. The results support the chosen approach, which relies on a small estimation data set and emphasizes importance of forecasts for market fundamentals. It is unlikely that statistical models with constant parameters would yield satisfying results.

Experiments with historical market data indicated that the framework was able to reproduce reasonable hourly price profiles, when the realized PDP volume was used as the demand. The results were highly sensitive to the estimated supply functions. Another challenge arose from cases, where the level of PDP demand in the estimation period was significantly different compared to the forecasting period. As a result, the estimated supply function was not defined for extreme values of PDP demand in the forecasting period. Automatic and/or statistical extrapolation of the estimated supply function turned out to be difficult, because its shape depends on fundamental factors. This could be necessary, for instance, if the temperature changes drastically.

Based on the results, it can be suggested that model user's input is incorporated with the estimation of supply functions. It may have several advantages when the framework is used in day-to-day forecasting. An experienced user is likely to surpass the estimation algorithm in accuracy. Instead of relying only on statistical data, the user can use their own view and interpret any information that could potentially have an effect on the pricing. The regression models used to explain changes in supply functions with fundamental changes did not produce reliable quantitative results. However, the trajectories of realized prices and PDP volumes can be used to study day-on-day changes in the pricing of supply. Combined with access to fundamental data, the analysis of trajectories can be useful in building market insight.

The scope of this thesis was limited to the Nordic *system price*, which is calculated without considering the limitations of transmission capacity between bidding areas. When a transmission line becomes congested, the concerned bidding areas will have different *area prices*. Other limitations arise from the assumptions made in the framework. It should be noted that the framework assumes that the supply function of PDP is constant within a day. The empirical studies revealed that the assumption does not hold all the times. Therefore, it is advised that the model user examines the data used for estimation in order to assess the validity of the assumption. Moreover, it is probable that the accuracy of empirical results suffered from the high-level aggregation of production volume data. Based on the available data, production volumes had to be classified as PIP or PDP by generation technology and country. The data did not differentiate regulated and unregulated hydro power generation. As a result, all Swedish and Finnish hydro power was considered to be PIP, even though it has some room for price-based regulation. Norwegian hydro power was considered to be fully PDP.

Further studies could concentrate on extending the proposed forecasting framework. A logical follow-up would be to model the area prices. As a starting point, the current framework could be applied to a single price area by modelling the area-specific PDP supply function, demand and exchange. Another improvement could be to treat the exchange in/out of the Nordic market (or between bidding areas) as an endogenous variable. At present exchange is assumed to be exogenous, while its direction and volume actually depend on the differences of area prices in the concerned areas. Furthermore, the supply function estimation technique has room for improvement. The heuristic algorithm presented in this thesis uses only historical price and volume data. It could be possible to attain more accurate supply functions by adjusting the estimate based on production tightness or other market fundamentals. More attention should be also given to modelling changes in the supply function as a function of market fundamentals. In this thesis, the analysis of fundamentals was based on realized historical values. If appropriate data could be obtained, it would be interesting to study if supply curve changes could be explained better by changes in consecutive fundamental forecasts. It seems possible, because water values and hydro power production schedules are based on future expectations.

Bibliography

Andersson B., Bergman L. (1995). *Market structure and the price of electricity: An ex ante analysis of the deregulated Swedish electricity market.* The Energy Journal **16**(2) pp. 97–109.

Angelus A. (2001). *Electricity price forecasting in deregulated markets.* The Electricity Journal **14**(3) pp. 32–41.

Baldick R., Grant R., Kahn E. (2004). *Theory and application of linear supply function equilibrium in electricity markets.* Journal of Regulatory Economics **25**(2) pp. 143–167.

Borenstein S., Bushnell J., Knittel C.R. (1999). *Market power in electricity markets: Beyond concentration measures.* The Energy Journal **20**(4) pp. 65–88.

Botterud A., Kristiansen T., Ilic M.D. (2010). *The relationship between spot and futures prices in the Nord Pool electricity market.* Energy Economics **32**(5) pp. 967–978.

Box G.E.P., Jenkins G.M. (1976). *Time Series Analysis: Forecasting and Control.* Holden-Day, Oakland, California.

Bunn D.W. (2000). *Forecasting loads and prices in competitive power markets.* Proceedings of the IEEE **88**(2) pp. 163–169.

Bunn D.W., Karakatsani N.V. (2003). *Forecasting electricity prices.* Working paper, London Business School. Also available as http://www.erasmusenergy.com/downloadattachment.php?aId=f1b845f3419ea055a3fe6fe11ac95ca8&articleId=50.

Bunn D.W., Oliveira F. (2001). *Agent-based simulation-an application to the new electricity trading arrangements of England and Wales.* IEEE Transactions on Evolutionary Computation **5**(5) pp. 493–503.

Bye T., Bruvoll A., Aune F.R. (2006). *The importance of volatility in inflow in a deregulated hydro-dominated power market.* Discussion papers no. 472, Statistics Norway, Research Department. Also available as http://www.ssb.no/cgi-bin/publsoek?job=forside&id=dp-472&kode=dp&lang=en.

Bye T., Hansen P.V. (2008). *How do spot prices affect aggregate electricity demand?* Discussion papers no. 527, Statistics Norway, Research Department.

Catalão J., Mariano S., Mendes V., Ferreira L. (2007). *Short-term electricity prices forecasting in a competitive market: A neural network approach*. Electric Power Systems Research **77**(10) pp. 1297–1304.

Dueholm L., Ravn H.F. (2004). *Modelling of short term electricity prices, hydro inflow and water values in the Norwegian hydro system*. In *6th IAEE European Conference*. Also available as http://www.balmorel.dk/Doc/ShortTermHydro2004.pdf.

European Market Coupling Company (2009). *EMCC Optimizer 2.0*. Technical report. Also available as http://www.marketcoupling.com/document/1137/EMCC%20Optimizer.pdf.

Eurostat (2012). *Supply, transformation, consumption – electricity – annual data [nrg_105a]*. Statistical database. Available as http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nrg_105a.

Eydeland A., Wolyniec K. (2003). *Energy and Power Risk Management*. Wiley.

Flatabø N., Doorman G., Grande O., Randen H., Wangensteen I. (2003). *Experience with the Nord Pool design and implementation*. IEEE Transactions on Power Systems **18**(2) pp. 541–547.

Fosso O., Gjelsvik A., Haugstad A., Mo B., Wangensteen I. (1999). *Generation scheduling in a deregulated system. The Norwegian case*. IEEE Transactions on Power Systems **14**(1) pp. 75–81.

Fridolfsson S.O., Tangerås T.P. (2009). *Market power in the Nordic electricity wholesale market: A survey of the empirical evidence*. Energy Policy **37**(9) pp. 3681–3692.

Førsund F.R. (2007). *Hydropower Economics*. Springer.

Gao F., Xiaohong G., Cao X.R., Papalexopoulos A. (2000). *Forecasting power market clearing price and quantity using a neural network method*. In *2000 Power Engineering Society Summer Meeting*, pp. 2183–2188, IEEE.

Garcia R., Contreras J., van Akkeren M., Garcia J. (2005). *A GARCH forecasting model to predict day-ahead electricity prices*. IEEE Transactions on Power Systems **20**(2) pp. 867–874.

Geman H., Roncoroni A. (2006). *Understanding the fine structure of electricity prices*. The Journal of Business **79**(3) pp. 1225–1261.

Gibbons R. (1992). *A Primer in Game Theory*. Prentice Hall.

Green R.J., Newbery D.M. (1992). *Competition in the British electricity spot market*. Journal of Political Economy **100**(5) pp. 929–953.

Guthrie G., Videbeck S. (2007). *Electricity spot price dynamics: Beyond financial models*. Energy Policy **35**(11) pp. 5614–5621.

Haykin S. (1994). *Neural Networks: A Comprehensive Foundation.* Prentice Hall International.

Huisman R., Huurman C., Mahieu R. (2007). *Hourly electricity prices in day-ahead markets.* Energy Economics **29**(2) pp. 240–248.

Jabłonska M. (2011). *From fluid dynamics to human psychology. What drives financial markets towards extreme events.* Ph.D. thesis, Lappeenranta University of Technology.

Javanainen T. (2005). *Analysis of short-term hydro power production in the Nordic electricity market.* Master's thesis, Helsinki University of Technology.

Jónsson T. (2008). *Forecasting of electricity prices: Accounting for wind power predictions.* Master's thesis, Technical University of Denmark.

Johnsen T.A. (2001). *Demand, generation and price in the Norwegian market for electric power.* Energy Economics **23**(3) pp. 227–251.

Kalatie S. (2005). *Sähkön markkinahintaa selittävät tekijät.* Master's thesis, Lappeenranta University of Technology.

Kanamura T., Ōhashi K. (2007). *A structural model for electricity prices with spikes: Measurement of spike risk and optimal policies for hydropower plant operation.* Energy Economics **29**(5) pp. 1010–1032.

Kara M. (editor) (2004). *Energia Suomessa: Tekniikka, talous ja ympäristövaikutukset.* VTT Prosessit, Edita Prima Oy.

Karakatsani N.V., Bunn D.W. (2008a). *Forecasting electricity prices: The impact of fundamentals and time-varying coefficients.* International Journal of Forecasting **24**(4) pp. 764–785.

Karakatsani N.V., Bunn D.W. (2008b). *Intra-day and regime-switching dynamics in electricity price formation.* Energy Economics **30**(4) pp. 1776–1797.

Kauppi O. (2009). *A model of imperfect dynamic competition.* Ph.D. thesis, Helsinki School of Economics.

Kumar David A., Wen F. (2001). *Market power in electricity supply.* IEEE Transactions on Energy Conversion **16**(4) pp. 352–360.

Laukkanen A. (2004). *Short term inflow forecasting in the Nordic power market.* Master's thesis, Helsinki University of Technology.

Livanis E., Zapranis A. (2007). *Forecasting the day-ahead electricity price in Nord Pool with neural networks: Some preliminary results.* Technical report, Energy Information Administration, Office of Integrated Analysis and Forecasting, U.S. Department of Energy. Also available as http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1568939.

Meyer-Brandis T., Tankov P. (2008). *Multi-factor jump-diffusion models of electricity prices.* International Journal of Theoretical and Applied Finance **11**(05) pp. 503–528.

Nogales F., Contreras J., Conejo A., Espinola R. (2002). *Forecasting next-day electricity prices by time series models.* IEEE Transactions on Power Systems **17**(2) pp. 342–348.

Nord Pool Spot (2011a). *Company presentation.* Accessed on 15.3.2011. http://nordpoolspot.com/Global/Download%20Center/Annual-report/Nord-Pool-Spot-Company-Presentation_2011.pdf

Nord Pool Spot (2011b). *The Nordic electricity exchange and the Nordic model for a liberalized electricity market.* Accessed on 4.10.2011. http://nordpoolspot.com/Global/Download%20Center/Rules-and-regulations/The-Nordic-Electricity-Exchange-and-the-Nordic-model-for-a-liberalized-electricity-market.pdf

Nord Pool Spot (2011c). *Trading appendix 2a: Elspot market regulations.* Accessed on 4.6.2012. http://nordpoolspot.com/Global/Download%20Center/Rules-and-regulations/trading-appendix_2-a_elspot-market-regulations_01jan2011.pdf

NordREG (2012). *Nordic market report 2012: Development in the Nordic electricity market.* Technical report. Also available as https://www.nordicenergyregulators.org/Publications/.

Pindyck R.S. (1999). *The long-run evolution of energy prices.* The Energy Journal **20**(2) pp. 1–27.

Queiroz A., Oliveira F., Lima J.M., Balestrassi P. (2007). *Simulating electricity spot prices in Brazil using neural network and design of experiments.* In *2007 IEEE Lausanne Power Tech*, pp. 2029–2034.

Ravn H.F. (2001). *Balmorel: A model for analyses of the electricity and CHP markets in the Baltic sea region.* Technical report, Balmorel Project. Also available as http://www.eabalmorel.dk/index.php?option=com_content&view=article&id=32:balmorel-a-model-for-analyses-of-the-electricity-and-chp-markets-&catid=2:pub&Itemid=22.

Seebregts A.J., Goldstein G.A., Smekens K. (2001). *Energy/Environmental modeling with the MARKAL family of models.* In *OR2001 Conference, Energy and Environment Session.*

Skantze P., Gubina A., Ilic M.D. (2000). *Bid-based stochastic model for electricity prices: The impact of fundamental drivers on market dynamics.* Energy Laboratory Publications MIT EL 00–004, Massachusetts Institute of Technology.

Szkuta B., Sanabria L., Dillon T. (1999). *Electricity price short-term forecasting* *using artificial neural networks*. IEEE Transactions on Power Systems **14**(3) pp. 851–857.

Vehviläinen I., Bröckl M., Virtanen E., Kallio M. (2010). *Nordic electricity peak prices during the winter 2009-2010*. Technical report, Gaia Consulting Oy. Also available as http://www.energiamarkkinavirasto.fi/files/Gaia_Raportti_20100930.pdf.

Vehviläinen I., Pyykkönen T. (2005). *Stochastic factor model for electricity spot price – the case of the Nordic market*. Energy Economics **27**(2) pp. 351–367.

Vuorinen A. (2009). *Planning of Optimal Power Systems: 2009 Edition*. Ekoenergo Oy.

Wallace S., Fleten S. (2003). *Stochastic programming models in energy*. Handbooks in Operations Research and Management Science **10** pp. 637–678.

Weron R. (2005). *Heavy tails and electricity prices*. In *The Deutsche Bundesbank's 2005 Annual Fall Conference*, pp. 10–12.

Weron R. (2007). *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Wiley.

Weron R., Misiorek A. (2006). *Short-term electricity price forecasting with time series models: A review and evaluation*. In *Complex Electricity Markets* (edited by W. Mielczarski), SEP Łódź.

Weron R., Misiorek A. (2008). *Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models*. International Journal of Forecasting **24**(4) pp. 744–763.

Wolfgang O., Haugstad A., Mo B., Gjelsvik A., Wangensteen I., Doorman G. (2009). *Hydro reservoir handling in Norway before and after deregulation*. Energy **34**(10) pp. 1642–1651.

World Wind Energy Association (2011). *World wind energy report 2010*. Technical report. Also available as http://www.wwindea.org/home/images/stories/pdfs/worldwindenergyreport2010_s.pdf.